

(19) World Intellectual Property Organization International Bureau





(43) International Publication Date 9 August 2001 (09.08.2001)

PCT

(10) International Publication Number WO 01/57188 A2

(51) International Patent Classification⁷:

- ~ -

Jose, CA 95117 (US). DRMANAC, Radoje, T. [YU/US];

Glovsky, and Popco, P.C., One Financial Center, Boston,

- (21) International Application Number: PCT/US01/03800
- (22) International Filing Date: 5 February 2001 (05.02.2001)
- (25) Filing Language:

English

C12N

(26) Publication Language:

English

(30) Priority Data:

09/496,914 3 February 2000 (03.02.2000) US 09/560,875 27 April 2000 (27.04.2000) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US 09/496,914 (CIP)
Filed on 3 February 2000 (03.02.2000)
US 09/560,875 (CIP)
Filed on 27 April 2000 (27.04.2000)

- (71) Applicant (for all designated States except US): HYSEQ, INC. [US/US]; 670 Almanor Avenue, Sunnyvalle, CA 94086 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): TANG, Y., Tom [US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). LIU, Chenghua [CN/US]; 1125 Ranchero Way, #14, San

850 East Greenwich Place, Palo Alto, CA 94303 (US).

(74) Agent: ELRIFI, Ivor, R.; Mintz, Levin, Coh. Ferris.

MA 02111 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM,

TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

 without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

1/57188 A2

(54) Title: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

(57) Abstract: The present invention provides novel nucleic acids, novel polypeptide sequences encoded by these nucleic acids and uses thereof.

NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

1. TECHNICAL FIELD

The present invention provides novel polynucleotides and proteins encoded by such polynucleotides, along with uses for these polynucleotides and proteins, for example in therapeutic, diagnostic and research methods.

2. BACKGROUND

5

25

30

35

Ignormalists and interferons and interferons and interferons and interferons are lymphokines, interferons, CSFs, chemokines, and interferons and interferons are lymphokines, interferons, CSFs, chemokines, and interferons and expression cloning techniques clone novel decade. The now routine hybridization cloning and expression cloning techniques clone novel polynucleotides "directly" in the sense that they rely on information directly related to the discovered protein (i.e., partial DNA/amino acid sequence of the protein in the case of hybridization cloning; activity of the protein in the case of expression cloning). More recent "indirect" cloning techniques such as signal sequence cloning, which isolates DNA sequences based on the presence of a now well-recognized secretory leader sequence motif, as well as various PCR-based or low stringency hybridization-based cloning techniques, have advanced the state of the art by making available large numbers of DNA/amino acid sequences for proteins that are known to have biological activity, for example, by virtue of their secreted nature in the case of leader sequence cloning, by virtue of their cell or tissue source in the case of PCR-based techniques, or by virtue of structural similarity to other genes of known biological activity.

Identified polynucleotide and polypeptide sequences have numerous applications in, for example, diagnostics, forensics, gene mapping; identification of mutations responsible for genetic disorders or other traits, to assess biodiversity, and to produce many other types of data and products dependent on DNA and amino acid sequences.

3. SUMMARY OF THE INVENTION

The compositions of the present invention include novel isolated polypeptides, novel isolated polynucleotides encoding such polypeptides, including recombinant DNA molecules, cloned genes or degenerate variants thereof, especially naturally occurring variants such as allelic variants, antisense polynucleotide molecules, and antibodies that specifically recognize one or more epitopes present on such polypeptides, as well as hybridomas producing such antibodies.

The compositions of the present invention additionally include vectors, including expression vectors, containing the polynucleotides of the invention, cells genetically engineered to contain such polynucleotides and cells genetically engineered to express such polynucleotides.

The present invention relates to a collection or library of at least one novel nucleic acid sequence assembled from expressed sequence tags (ESTs) isolated mainly by sequencing by hybridization (SBH), and in some cases, sequences obtained from one or more public databases. The invention relates also to the proteins encoded by such polynucleotides, along with therapeutic, diagnostic and research utilities for these polynucleotides and proteins. These nucleic acid sequences are designated as SEQ ID NO: 1-1350. The polypeptides sequences are designated SEQ ID NO: 1351-2700. The nucleic acids and polypeptides are provided in the Sequence Listing. In the nucleic acids provided in the Sequence Listing, A is adenosine; C is cytosine; G is guanine; T is thymine; and N is any of the four bases. In the amino acids provided in the Sequence Listing, * corresponds to the stop codon.

5

10

15

20

25

30

The nucleic acid sequences of the present invention also include, nucleic acid sequences that hybridize to the complement of SEQ ID NO:1-1350 under stringent hybridization conditions; nucleic acid sequences which are allelic variants or species homologues of any of the nucleic acid sequences recited above, or nucleic acid sequences that encode a peptide comprising a specific domain or truncation of the peptides encoded by SEQ ID NO:1-1350. A polynucleotide comprising a nucleotide sequence having at least 90% identity to an identifying sequence of SEQ ID NO:1-1350 or a degenerate variant or fragment thereof. The identifying sequence can be 100 base pairs in length.

The nucleic acid sequences of the present invention also include the sequence information from the nucleic acid sequences of SEQ ID NO:1-1350. The sequence information can be a segment of any one of SEQ ID NO:1-1350 that uniquely identifies or represents the sequence information of SEQ ID NO:1-1350.

A collection as used in this application can be a collection of only one polynucleotide. The collection of sequence information or identifying information of each sequence can be provided on a nucleic acid array. In one embodiment, segments of sequence information is provided on a nucleic acid array to detect the polynucleotide that contains the segment. The array can be designed to detect full-match or mismatch to the polynucleotide that contains the segment. The collection can also be provided in a computer-readable format.

This invention also includes the reverse or direct complement of any of the nucleic acid sequences recited above; cloning or expression vectors containing the nucleic acid sequences; and host cells or organisms transformed with these expression vectors. Nucleic acid sequences (or their reverse or direct complements) according to the invention have numerous applications in a variety of techniques known to those skilled in the art of molecular biology, such as use as hybridization probes, use as primers for PCR, use in an array, use in computer-readable media, use in sequencing

full-length genes, use for chromosome and gene mapping, use in the recombinant production of protein, and use in the generation of anti-sense DNA or RNA, their chemical analogs and the like.

In a preferred embodiment, the nucleic acid sequences of SEQ ID NO:1-1350 or novel segments or parts of the nucleic acids of the invention are used as primers in expression assays that are well known in the art. In a particularly preferred embodiment, the nucleic acid sequences of SEQ ID NO:1-1350 or novel segments or parts of the nucleic acids provided herein are used in diagnostics for identifying expressed genes or, as well known in the art and exemplified by Vollrath et al., Science 258:52-59 (1992), as expressed sequence tags for physical mapping of the human genome.

5

10

15

20

25

30

The isolated polynucleotides of the invention include, but are not limited to, a polynucleotide comprising any one of the nucleotide sequences set forth in SEQ ID NO:1-1350; a polynucleotide comprising any of the full length protein coding sequences of SEQ ID NO:1 - 1350; and a polynucleotide comprising any of the nucleotide sequences of the mature protein coding sequences of SEQ ID NO: 1-1350. The polynucleotides of the present invention also include, but are not limited to, a polynucleotide that hybridizes under stringent hybridization conditions to (a) the complement of any one of the nucleotide sequences set forth in SEQ ID NO:1-1350; (b) a nucleotide sequence encoding any one of the amino acid sequences set forth in the Sequence Listing (e.g., SEQ ID NO: 1351-2700); (c) a polynucleotide which is an allelic variant of any polynucleotides recited above; (d) a polynucleotide which encodes a species homolog (e.g. orthologs) of any of the proteins recited above; or (e) a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of any of the polypeptides comprising an amino acid sequence set forth in the Sequence Listing.

The isolated polypeptides of the invention include, but are not limited to, a polypeptide comprising any of the amino acid sequences set forth in the Sequence Listing; or the corresponding full length or mature protein. Polypeptides of the invention also include polypeptides with biological activity that are encoded by (a) any of the polynucleotides having a nucleotide sequence set forth in SEQ ID NO:1-1350; or (b) polynucleotides that hybridize to the complement of the polynucleotides of (a) under stringent hybridization conditions. Biologically or immunologically active variants of any of the polypeptide sequences in the Sequence Listing, and "substantial equivalents" thereof (e.g., with at least about 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98% or 99% amino acid sequence identity) that preferably retain biological activity are also contemplated. The polypeptides of the invention may be wholly or partially chemically synthesized but are preferably produced by recombinant means using the genetically engineered cells (e.g. host cells) of the invention.

The invention also provides compositions comprising a polypeptide of the invention. Polypeptide compositions of the invention may further comprise an acceptable carrier, such as a hydrophilic, *e.g.*, pharmaceutically acceptable, carrier.

The invention also provides host cells transformed or transfected with a polynucleotide of the invention.

5

10

15

20

25

30

35

The invention also relates to methods for producing a polypeptide of the invention comprising growing a culture of the host cells of the invention in a suitable culture medium under conditions permitting expression of the desired polypeptide, and purifying the polypeptide from the culture or from the host cells. Preferred embodiments include those in which the protein produced by such process is a mature form of the protein.

Polynucleotides according to the invention have numerous applications in a variety of techniques known to those skilled in the art of molecular biology. These techniques include use as hybridization probes, use as oligomers, or primers, for PCR, use for chromosome and gene mapping, use in the recombinant production of protein, and use in generation of anti-sense DNA or RNA, their chemical analogs and the like. For example, when the expression of an mRNA is largely restricted to a particular cell or tissue type, polynucleotides of the invention can be used as hybridization probes to detect the presence of the particular cell or tissue mRNA in a sample using, *e.g.*, in situ hybridization.

In other exemplary embodiments, the polynucleotides are used in diagnostics as expressed sequence tags for identifying expressed genes or, as well known in the art and exemplified by Vollrath et al., Science 258:52-59 (1992), as expressed sequence tags for physical mapping of the human genome.

The polypeptides according to the invention can be used in a variety of conventional procedures and methods that are currently applied to other proteins. For example, a polypeptide of the invention can be used to generate an antibody that specifically binds the polypeptide. Such antibodies, particularly monoclonal antibodies, are useful for detecting or quantitating the polypeptide in tissue. The polypeptides of the invention can also be used as molecular weight markers, and as a food supplement.

Methods are also provided for preventing, treating, or ameliorating a medical condition which comprises the step of administering to a mammalian subject a therapeutically effective amount of a composition comprising a polypeptide of the present invention and a pharmaceutically acceptable carrier.

In particular, the polypeptides and polynucleotides of the invention can be utilized, for example, in methods for the prevention and/or treatment of disorders involving aberrant protein expression or biological activity.

The present invention further relates to methods for detecting the presence of the polynucleotides or polypeptides of the invention in a sample. Such methods can, for example, butilized as part of prognostic and diagnostic evaluation of disorders as recited herein and for the identification of subjects exhibiting a predisposition to such conditions. The invention provides a method for detecting the polynucleotides of the invention in a sample, comprising contacting the sample with a compound that binds to and forms a complex with the polynucleotide of interest for a period sufficient to form the complex and under conditions sufficient to form a complex and detecting the complex such that if a complex is detected, the polynucleotide of interest is detected. The invention also provides a method for detecting the polypeptides of the invention in a sample comprising contacting the sample with a compound that binds to and form a complex with the polypeptide under conditions and for a period sufficient to form the complex and detecting the formation of the complex such that if a complex is formed, the polypeptide is detected.

10

15

20

25

30

35

The invention also provides kits comprising polynucleotide probes and/or monoclonal antibodies, and optionally quantitative standards, for carrying out methods of the invention. Furthermore, the invention provides methods for evaluating the efficacy of drugs, and monitoring the progress of patients, involved in clinical trials for the treatment of disorders as recited above.

The invention also provides methods for the identification of compounds that modulate (i.e., increase or decrease) the expression or activity of the polynucleotides and/or polypeptides of the invention. Such methods can be utilized, for example, for the identification of compounds that can ameliorate symptoms of disorders as recited herein. Such methods can include, but are not limited to, assays for identifying compounds and other substances that interact with (e.g., bind to) the polypeptides of the invention. The invention provides a method for identifying a compound that binds to the polypeptides of the invention comprising contacting the compound with a polypeptide of the invention in a cell for a time sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a reporter gene sequence expression such that if expression of the reporter gene is detected the compound the binds to a polypeptide of the invention is identified.

The methods of the invention also provides methods for treatment which involve the administration of the polynucleotides or polypeptides of the invention to individuals exhibiting symptoms or tendencies. In addition, the invention encompasses methods for treating diseases o disorders as recited herein comprising administering compounds and other substances that modulate the overall activity of the target gene products. Compounds and other substances can

effect such modulation either on the level of target gene/protein expression or target protein activity.

The polypeptides of the present invention and the polynucleotides encoding them are also useful for the same functions known to one of skill in the art as the polypeptides and polynucleotides to which they have homology (set forth in Table 2). If no homology is set forth for a sequence, then the polypeptides and polynucleotides of the present invention are useful for a variety of applications, as described herein, including use in arrays for detection.

4. DETAILED DESCRIPTION OF THE INVENTION

4.1 DEFINITIONS

5

10

15

20

25

30

35

It must be noted that as used herein and in the appended claims, the singular forms "a", "an" and "the" include plural references unless the context clearly dictates otherwise.

The term "active" refers to those forms of the polypeptide which retain the biologic and/or immunologic activities of any naturally occurring polypeptide. According to the invention, the terms "biologically active" or "biological activity" refer to a protein or peptide having structural, regulatory or biochemical functions of a naturally occurring molecule. Likewise "immunologically active" or "immunological activity" refers to the capability of the natural, recombinant or synthetic polypeptide to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The term "activated cells" as used in this application are those cells which are engaged in extracellular or intracellular membrane trafficking, including the export of secretory or enzymatic molecules as part of a normal or disease process.

The terms "complementary" or "complementarity" refer to the natural binding of polynucleotides by base pairing. For example, the sequence 5'-AGT-3' binds to the complementary sequence 3'-TCA-5'. Complementarity between two single-stranded molecules may be "partial" such that only some of the nucleic acids bind or it may be "complete" such that total complementarity exists between the single stranded molecules. The degree of complementarity between the nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands.

The term "embryonic stem cells (ES)" refers to a cell that can give rise to many differentiated cell types in an embryo or an adult, including the germ cells. The term "germ line stem cells (GSCs)" refers to stem cells derived from primordial stem cells that provide a steady and continuous source of germ cells for the production of gametes. The term "primordial germ

cells (PGCs)" refers to a small population of cells set aside from other cell lineages particularly from the yolk sac, mesenteries, or gonadal ridges during embryogenesis that have the potential to differentiate into germ cells and other cells. PGCs are the source from which GSCs and ES cells are derived. The PGCs, the GSCs and the ES cells are capable of self-renewal. Thus these cells not only populate the germ line and give rise to a plurality of terminally differentiated cells that comprise the adult specialized organs, but are able to regenerate themselves.

The term "expression modulating fragment," EMF, means a series of nucleotides which modulates the expression of an operably linked ORF or another EMF.

5

10

15

20

25

30

35

As used herein, a sequence is said to "modulate the expression of an operably linked sequence" when the expression of the sequence is altered by the presence of the EMF. EMFs include, but are not limited to, promoters, and promoter modulating sequences (inducible elements). One class of EMFs are nucleic acid fragments which induce the expression of an operably linked ORF in response to a specific regulatory factor or physiological event.

The terms "nucleotide sequence" or "nucleic acid" or "polynucleotide" or "oligonculeotide" are used interchangeably and refer to a heteropolymer of nucleotides or the sequence of these nucleotides. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA) or to any DNA-like or RNA-like material. In the sequences herein A is adenine, C is cytosine, T is thymine, G is guanine and N is A, C, G or T (U). It is contemplated that where the polynucleotide is RNA, the T (thymine) in the sequences provided herein is substituted with U (uracil). Generally, nucleic acid segments provided by this invention may be assembled from fragments of the genome and short oligonucleotide linkers, or from a series of oligonucleotides, or from individual nucleotides, to provide a synthetic nucleic acid which is capable of being expressed in a recombinant transcriptional unit comprising regulatory elements derived from a microbial or viral operon, or a eukaryotic gene.

The terms "oligonucleotide fragment" or a "polynucleotide fragment", "portion," or "segment" or "probe" or "primer" are used interchangeably and refer to a sequence of nucleotide residues which are at least about 5 nucleotides, more preferably at least about 7 nucleotides, more preferably at least about 11 nucleotides and most preferably at least about 17 nucleotides. The fragment is preferably less than about 500 nucleotides, preferably less than about 200 nucleotides, more preferably less than about 100 nucleotides, more preferably less than about 50 nucleotides, more preferably less than 30 nucleotides. Preferably the probe is from about 6 nucleotides to about 200 nucleotides, preferably from about 15 to about 50 nucleotides, more preferably from about 17 to 30 nucleotides and most preferably from about 20 to 25 nucleotides. Preferably the fragments can

be used in polymerase chain reaction (PCR), various hybridization procedures or microarray procedures to identify or amplify identical or related parts of mRNA or DNA molecules. A fragment or segment may uniquely identify each polynucleotide sequence of the present invention. Preferably the fragment comprises a sequence substantially similar to any one of SEQ ID NOs:1-1350.

5

10

15

20

25

30

Probes may, for example, be used to determine whether specific mRNA molecules are present in a cell or tissue or to isolate similar nucleic acid sequences from chromosomal DNA as described by Walsh et al. (Walsh, P.S. et al., 1992, PCR Methods Appl 1:241-250). They may be labeled by nick translation, Klenow fill-in reaction, PCR, or other methods well known in the art. Probes of the present invention, their preparation and/or labeling are elaborated in Sambrook, J. et al., 1989, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY; or Ausubel, F.M. et al., 1989, Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, both of which are incorporated herein by reference in their entirety.

The nucleic acid sequences of the present invention also include the sequence information from the nucleic acid sequences of SEQ ID NO:1-1350. The sequence information can be a segment of any one of SEQ ID NO:1-1350 that uniquely identifies or represents the sequence information of that sequence of SEQ ID NO:1-1350. One such segment can be a twenty-mer nucleic acid sequence because the probability that a twenty-mer is fully matched in the human genome is 1 in 300. In the human genome, there are three billion base pairs in one set of chromosomes. Because 4²⁰ possible twenty-mers exist, there are 300 times more twenty-mers than there are base pairs in a set of human chromosomes. Using the same analysis, the probability for a seventeen-mer to be fully matched in the human genome is approximately 1 in 5. When these segments are used in arrays for expression studies, fifteen-mer segments can be used. The probability that the fifteen-mer is fully matched in the expressed sequences is also approximately one in five because expressed sequences comprise less than approximately 5% of the entire genome sequence.

Similarly, when using sequence information for detecting a single mismatch, a segment can be a twenty-five mer. The probability that the twenty-five mer would appear in a human genome with a single mismatch is calculated by multiplying the probability for a full match $(1 \div 4^{25})$ times the increased probability for mismatch at each nucleotide position (3×25) . The probability that an eighteen mer with a single mismatch can be detected in an array for expression studies is approximately one in five. The probability that a twenty-mer with a single mismatch can be detected in a human genome is approximately one in five.

The term "open reading frame," ORF, means a series of nucleotide triplets coding for amino acids without any termination codons and is a sequence translatable into protein.

The terms "operably linked" or "operably associated" refer to functionally related nucleic acid sequences. For example, a promoter is operably associated or operably linked with a coding sequence if the promoter controls the transcription of the coding sequence. While operably linked nucleic acid sequences can be contiguous and in the same reading frame, certain genetic elements e.g. repressor genes are not contiguously linked to the coding sequence but still control transcription/translation of the coding sequence.

5

10

15

20

25

30

The term "pluripotent" refers to the capability of a cell to differentiate into a number of differentiated cell types that are present in an adult organism. A pluripotent cell is restricted in its differentiation capability in comparison to a totipotent cell.

The terms "polypeptide" or "peptide" or "amino acid sequence" refer to an oligopeptide, peptide, polypeptide or protein sequence or fragment thereof and to naturally occurring or synthetic molecules. A polypeptide "fragment," "portion," or "segment" is a stretch of amino acid residues of at least about 5 amino acids, preferably at least about 7 amino acids, more preferably at least about 9 amino acids and most preferably at least about 17 or more amino acids. The peptide preferably is not greater than about 200 amino acids, more preferably less than 150 amino acids and most preferably less than 100 amino acids. Preferably the peptide is from about 5 to about 200 amino acids. To be active, any polypeptide must have sufficient length to display biological and/or immunological activity.

The term "naturally occurring polypeptide" refers to polypeptides produced by cells that have not been genetically engineered and specifically contemplates various polypeptides arising from post-translational modifications of the polypeptide including, but not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation and acylation.

The term "translated protein coding portion" means a sequence which encodes for the full length protein which may include any leader sequence or any processing sequence.

The term "mature protein coding sequence" means a sequence which encodes a peptide or protein without a signal or leader sequence. The "mature protein portion" means that portion of the protein which does not include a signal or leader sequence. The peptide may have been produced by processing in the cell which removes any leader/signal sequence. The mature protein portion may or may not include the initial methionine residue. The methionine residue may be removed from the protein during processing in the cell. The peptide may be produced synthetically or the protein may have been produced using a polynucleotide only encoding for the mature protein coding sequence.

The term "derivative" refers to polypeptides chemically modified by such techniques as ubiquitination, labeling (e.g., with radionuclides or various enzymes), covalent polymer attachment such as pegylation (derivatization with polyethylene glycol) and insertion or substitution by chemical synthesis of amino acids such as ornithine, which do not normally occur in human proteins.

5

10

15

20

25

30

35

The term "variant" (or "analog") refers to any polypeptide differing from naturally occurring polypeptides by amino acid insertions, deletions, and substitutions, created using, e.g., recombinant DNA techniques. Guidance in determining which amino acid residues may be replaced, added or deleted without abolishing activities of interest, may be found by comparing the sequence of the particular polypeptide with that of homologous peptides and minimizing the number of amino acid sequence changes made in regions of high homology (conserved regions) or by replacing amino acids with consensus sequence.

Alternatively, recombinant variants encoding these same or similar polypeptides may be synthesized or selected by making use of the "redundancy" in the genetic code. Various codon substitutions, such as the silent changes which produce various restriction sites, may be introduced to optimize cloning into a plasmid or viral vector or expression in a particular prokaryotic or eukaryotic system. Mutations in the polynucleotide sequence may be reflected in the polypeptide or domains of other peptides added to the polypeptide to modify the properties of any part of the polypeptide, to change characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate.

Preferably, amino acid "substitutions" are the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, *i.e.*, conservative amino acid replacements. "Conservative" amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved. For example, nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine; polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine; positively charged (basic) amino acids include arginine, lysine, and histidine; and negatively charged (acidic) amino acids include aspartic acid and glutamic acid. "Insertions" or "deletions" are preferably in the range of about 1 to 20 amino acids, more preferably 1 to 10 amino acids. The variation allowed may be experimentally determined by systematically making insertions, deletions, or substitutions of amino acids in a polypeptide molecule using recombinant DNA techniques and assaying the resulting recombinant variants for activity.

Alternatively, where alteration of function is desired, insertions, deletions or non-conservative alterations can be engineered to produce altered polypeptides. Such alterations

can, for example, alter one or more of the biological functions or biochemical characteristics of the polypeptides of the invention. For example, such alterations may change polypeptide characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate. Further, such alterations can be selected so as to generate polypeptides that are better suited for expression, scale up and the like in the host cells chosen for expression. For example, cysteine residues can be deleted or substituted with another amino acid residue in order to eliminate disulfide bridges.

5

10

15

20

25

30

35

The terms "purified" or "substantially purified" as used herein denotes that the indicated nucleic acid or polypeptide is present in the substantial absence of other biological macromolecules, e.g., polynucleotides, proteins, and the like. In one embodiment, the polynucleotide or polypeptide is purified such that it constitutes at least 95% by weight, more preferably at least 99% by weight, of the indicated biological macromolecules present (but water, buffers, and other small molecules, especially molecules having a molecular weight of less than 1000 daltons, can be present).

The term "isolated" as used herein refers to a nucleic acid or polypeptide separated from at least one other component (e.g., nucleic acid or polypeptide) present with the nucleic acid or polypeptide in its natural source. In one embodiment, the nucleic acid or polypeptide is found in the presence of (if anything) only a solvent, buffer, ion, or other component normally present in a solution of the same. The terms "isolated" and "purified" do not encompass nucleic acids or polypeptides present in their natural source.

The term "recombinant," when used herein to refer to a polypeptide or protein, means that a polypeptide or protein is derived from recombinant (e.g., microbial, insect, or mammalian) expression systems. "Microbial" refers to recombinant polypeptides or proteins made in bacterial or fungal (e.g., yeast) expression systems. As a product, "recombinant microbial" defines a polypeptide or protein essentially free of native endogenous substances and unaccompanied by associated native glycosylation. Polypeptides or proteins expressed in most bacterial cultures, e.g., E. coli, will be free of glycosylation modifications; polypeptides or proteins expressed in yeast will have a glycosylation pattern in general different from those expressed in mammalian cells.

The term "recombinant expression vehicle or vector" refers to a plasmid or phage or virus or vector, for expressing a polypeptide from a DNA (RNA) sequence. An expression vehicle can comprise a transcriptional unit comprising an assembly of (1) a genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers, (2) a structural or coding sequence which is transcribed into mRNA and translated into protein, and (3) appropriate transcription initiation and termination sequences. Structural units intended for use

in yeast or eukaryotic expression systems preferably include a leader sequence enabling extracellular secretion of translated protein by a host cell. Alternatively, where recombinant protein is expressed without a leader or transport sequence, it may include an amino terminal methionine residue. This residue may or may not be subsequently cleaved from the expressed recombinant protein to provide a final product.

5

10

15

20

25

30

35

The term "recombinant expression system" means host cells which have stably integrated a recombinant transcriptional unit into chromosomal DNA or carry the recombinant transcriptional unit extrachromosomally. Recombinant expression systems as defined herein will express heterologous polypeptides or proteins upon induction of the regulatory elements linked to the DNA segment or synthetic gene to be expressed. This term also means host cells which have stably integrated a recombinant genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers. Recombinant expression systems as defined herein will express polypeptides or proteins endogenous to the cell upon induction of the regulatory elements linked to the endogenous DNA segment or gene to be expressed. The cells can be prokaryotic or eukaryotic.

The term "secreted" includes a protein that is transported across or through a membrane, including transport as a result of signal sequences in its amino acid sequence when it is expressed in a suitable host cell. "Secreted" proteins include without limitation proteins secreted wholly (e.g., soluble proteins) or partially (e.g., receptors) from the cell in which they are expressed. "Secreted" proteins also include without limitation proteins that are transported across the membrane of the endoplasmic reticulum. "Secreted" proteins are also intended to include proteins containing non-typical signal sequences (e.g. Interleukin-1 Beta, see Krasney, P.A. and Young, P.R. (1992) Cytokine 4(2):134-143) and factors released from damaged cells (e.g. Interleukin-1 Receptor Antagonist, see Arend, W.P. et. al. (1998) Annu. Rev. Immunol. 16:27-55)

Where desired, an expression vector may be designed to contain a "signal or leader sequence" which will direct the polypeptide through the membrane of a cell. Such a sequence may be naturally present on the polypeptides of the present invention or provided from heterologous protein sources by recombinant DNA techniques.

The term "stringent" is used to refer to conditions that are commonly understood in the art as stringent. Stringent conditions can include highly stringent conditions (i.e., hybridization to filter-bound DNA in 0.5 M NaHPO₄, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°C, and washing in 0.1X SSC/0.1% SDS at 68°C), and moderately stringent conditions (i.e., washing in 0.2X SSC/0.1% SDS at 42°C). Other exemplary hybridization conditions are described herein in the examples.

In instances of hybridization of deoxyoligonucleotides, additional exemplary stringent hybridization conditions include washing in 6X SSC/0.05% sodium pyrophosphate at 37°C (for 14-base oligonucleotides), 48°C (for 17-base oligos), 55°C (for 20-base oligonucleotides), and 60°C (for 23-base oligonucleotides).

5

10

15

20

25

30

As used herein, "substantially equivalent" can refer both to nucleotide and amino acid sequences, for example a mutant sequence, that varies from a reference sequence by one or more substitutions, deletions, or additions, the net effect of which does not result in an adverse functional dissimilarity between the reference and subject sequences. Typically, such a substantially equivalent sequence varies from one of those listed herein by no more than about 35% (i.e., the number of individual residue substitutions, additions, and/or deletions in a substantially equivalent sequence, as compared to the corresponding reference sequence, divided by the total number of residues in the substantially equivalent sequence is about 0.35 or less). Such a sequence is said to have 65% sequence identity to the listed sequence. In one embodiment, a substantially equivalent, e.g., mutant, sequence of the invention varies from a listed sequence by no more than 30% (70% sequence identity); in a variation of this embodiment, by no more than 25% (75% sequence identity); and in a further variation of this embodiment, by no more than 20% (80% sequence identity) and in a further variation of this embodiment, by no more than 10% (90% sequence identity) and in a further variation of this embodiment, by no more that 5% (95% sequence identity). Substantially equivalent, e.g., mutant, amino acid sequences according to the invention preferably have at least 80% sequence identity with a listed amino acid sequence, more preferably at least 85% sequence identity, more preferably at least 90% sequence identity, more preferably at least 95% identity, more preferably at least 98% identity, and most preferably at least 99% identity. Substantially equivalent nucleotide sequences of the invention can have lower percent sequence identities, taking into account, for example, the redundancy or degeneracy of the genetic code. Preferably, nucleotide sequence has at least about 65% identity, more preferably at least about 75% identity, more preferably at least about 80% sequence identity, more preferably at least about 85% sequence identity, more preferably at least about 90% sequence identity, and most preferably at least about 95% identity, more preferably at least about 98% sequence identity, and most preferably at least about 99% sequence identity. For the purposes of the present invention, sequences having substantially equivalent biological activity and substantially equivalent expression characteristics are considered substantially equivalent. For the purposes of determining equivalence, truncation of the mature sequence (e.g., via a mutation which creates a spurious stop codon) should be disregarded. Sequence identity may be determined, e.g., using the Jotun Hein method (Hein, J.

(1990) Methods Enzymol. 183:626-645). Identity between sequences can also be determined by other methods known in the art, e.g. by varying hybridization conditions.

The term "totipotent" refers to the capability of a cell to differentiate into all of the cell types of an adult organism.

The term "transformation" means introducing DNA into a suitable host cell so that the DNA is replicable, either as an extrachromosomal element, or by chromosomal integration. The term "transfection" refers to the taking up of an expression vector by a suitable host cell, whether or not any coding sequences are in fact expressed. The term "infection" refers to the introduction of nucleic acids into a suitable host cell by use of a virus or viral vector.

As used herein, an "uptake modulating fragment," UMF, means a series of nucleotides which mediate the uptake of a linked DNA fragment into a cell. UMFs can be readily identified using known UMFs as a target sequence or target motif with the computer-based systems described below. The presence and activity of a UMF can be confirmed by attaching the suspected UMF to a marker sequence. The resulting nucleic acid molecule is then incubated with an appropriate host under appropriate conditions and the uptake of the marker sequence is determined. As described above, a UMF will increase the frequency of uptake of a linked marker sequence.

Each of the above terms is meant to encompass all that is described for each, unless the context dictates otherwise.

4.2 NUCLEIC ACIDS OF THE INVENTION

5

10

15

20

25

30

35

Nucleotide sequences of the invention are set forth in the Sequence Listing.

The isolated polynucleotides of the invention include a polynucleotide comprising the nucleotide sequences of SEQ ID NO:1-1350; a polynucleotide encoding any one of the peptide sequences of SEQ ID NO:1351-2700; and a polynucleotide comprising the nucleotide sequence encoding the mature protein coding sequence of the polypeptides of any one of SEQ ID NO:1351-2700. The polynucleotides of the present invention also include, but are not limited to, a polynucleotide that hybridizes under stringent conditions to (a) the complement of any of the nucleotides sequences of SEQ ID NO:1-1350; (b) nucleotide sequences encoding any one of the amino acid sequences set forth in the Sequence Listing; (c) a polynucleotide which is an allelic variant of any polynucleotide recited above; (d) a polynucleotide which encodes a species homolog of any of the proteins recited above; or (e) a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of the polypeptides of SEQ ID NO: 1351-2700. Domains of interest may depend on the nature of the encoded polypeptide; e.g., domains in receptor-like polypeptides include ligand-binding, extracellular, transmembrane, or cytoplasmic

domains, or combinations thereof; domains in immunoglobulin-like proteins include the variable immunoglobulin-like domains; domains in enzyme-like polypeptides include catalytic and substrate binding domains; and domains in ligand polypeptides include receptor-binding domains.

The polynucleotides of the invention include naturally occurring or wholly or partially synthetic DNA, e.g., cDNA and genomic DNA, and RNA, e.g., mRNA. The polynucleotides may include all of the coding region of the cDNA or may represent a portion of the coding region of the cDNA.

5

10

15

20

25

30

35

The present invention also provides genes corresponding to the cDNA sequences disclosed herein. The corresponding genes can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include the preparation of probes or primers from the disclosed sequence information for identification and/or amplification of genes in appropriate genomic libraries or other sources of genomic materials. Further 5' and 3' sequence can be obtained using methods known in the art. For example, full length cDNA or genomic DNA that corresponds to any of the polynucleotides of SEQ ID NO:1-1350 can be obtained by screening appropriate cDNA or genomic DNA libraries under suitable hybridization conditions using any of the polynucleotides of SEQ ID NO:1-1350 or a portion thereof as a probe. Alternatively, the polynucleotides of SEQ ID NO:1-1350 may be used as the basis for suitable primer(s) that allow identification and/or amplification of genes in appropriate genomic DNA or cDNA libraries.

The nucleic acid sequences of the invention can be assembled from ESTs and sequences (including cDNA and genomic sequences) obtained from one or more public databases, such as dbEST, gbpri, and UniGene. The EST sequences can provide identifying sequence information, representative fragment or segment information, or novel segment information for the full-length gene.

The polynucleotides of the invention also provide polynucleotides including nucleotide sequences that are substantially equivalent to the polynucleotides recited above. Polynucleotides according to the invention can have, e.g., at least about 65%, at least about 70%, at least about 75%, at least about 80%, 81%, 82%, 83%, 84%, more typically at least about 85%, 86%, 87%, 88%, 89%, more typically at least about 90%, 91%, 92%, 93%, 94%, and even more typically at least about 95%, 96%, 97%, 98%, 99%, sequence identity to a polynucleotide recited above.

Included within the scope of the nucleic acid sequences of the invention are nucleic acid sequence fragments that hybridize under stringent conditions to any of the nucleotide sequences of SEQ ID NO:1-1350, or complements thereof, which fragment is greater than about 5 nucleotides, preferably 7 nucleotides, more preferably greater than 9 nucleotides and most preferably greater than 17 nucleotides. Fragments of, e.g. 15, 17, or 20 nucleotides or more that

are selective for (*i.e.* specifically hybridize to any one of the polynucleotides of the invention) are contemplated. Probes capable of specifically hybridizing to a polynucleotide can differentiate polynucleotide sequences of the invention from other polynucleotide sequences in the same family of genes or can differentiate human genes from genes of other species, and are preferably based on unique nucleotide sequences.

5

10

15

20

25

30

35

The sequences falling within the scope of the present invention are not limited to these specific sequences, but also include allelic and species variations thereof. Allelic and species variations can be routinely determined by comparing the sequence provided SEQ ID NO:1-1350, a representative fragment thereof, or a nucleotide sequence at least 90% identical, preferably 95% identical, to SEQ ID NO:1-1350 with a sequence from another isolate of the same species. Furthermore, to accommodate codon variability, the invention includes nucleic acid molecules coding for the same amino acid sequences as do the specific ORFs disclosed herein. In other words, in the coding region of an ORF, substitution of one codon for another codon that encodes the same amino acid is expressly contemplated.

The nearest neighbor or homology result for the nucleic acids of the present invention, including SEQ ID NO:1-1350, can be obtained by searching a database using an algorithm or a program. Preferably, a BLAST which stands for Basic Local Alignment Search Tool is used to search for local sequence alignments (Altshul, S.F. J Mol. Evol. 36 290-300 (1993) and Altschul S.F. et al. J. Mol. Biol. 21:403-410 (1990)). Alternatively a FASTA version 3 search against Genpept, using Fastxy algorithm.

Species homologs (or orthologs) of the disclosed polynucleotides and proteins are also provided by the present invention. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from the desired species.

The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotide which also encode proteins which are identical, homologous or related to that encoded by the polynucleotides.

The nucleic acid sequences of the invention are further directed to sequences which encode variants of the described nucleic acids. These amino acid sequence variants may be prepared by methods known in the art by introducing appropriate nucleotide changes into a native or variant polynucleotide. There are two variables in the construction of amino acid sequence variants: the location of the mutation and the nature of the mutation. Nucleic acids encoding the amino acid sequence variants are preferably constructed by mutating the polynucleotide to encode an amino acid sequence that does not occur in nature. These nucleic

acid alterations can be made at sites that differ in the nucleic acids from different species (variable positions) or in highly conserved regions (constant regions). Sites at such locations will typically be modified in series, e.g., by substituting first with conservative choices (e.g., hydrophobic amino acid to a different hydrophobic amino acid) and then with more distant choices (e.g., hydrophobic amino acid to a charged amino acid), and then deletions or insertions may be made at the target site. Amino acid sequence deletions generally range from about 1 to 30 residues, preferably about 1 to 10 residues, and are typically contiguous. Amino acid insertions include amino- and/or carboxyl-terminal fusions ranging in length from one to one hundred or more residues, as well as intrasequence insertions of single or multiple amino acid residues. Intrasequence insertions may range generally from about 1 to 10 amino residues, preferably from 1 to 5 residues. Examples of terminal insertions include the heterologous signal sequences necessary for secretion or for intracellular targeting in different host cells and sequences such as FLAG or poly-histidine sequences useful for purifying the expressed protein.

In a preferred method, polynucleotides encoding the novel amino acid sequences are changed via site-directed mutagenesis. This method uses oligonucleotide sequences to alter a polynucleotide to encode the desired amino acid variant, as well as sufficient adjacent nucleotides on both sides of the changed amino acid to form a stable duplex on either side of the site of being changed. In general, the techniques of site-directed mutagenesis are well known to those of skill in the art and this technique is exemplified by publications such as, Edelman et al., *DNA* 2:183 (1983). A versatile and efficient method for producing site-specific changes in a polynucleotide sequence was published by Zoller and Smith, *Nucleic Acids Res.* 10:6487-6500 (1982). PCR may also be used to create amino acid sequence variants of the novel nucleic acids. When small amounts of template DNA are used as starting material, primer(s) that differs slightly in sequence from the corresponding region in the template DNA can generate the desired amino acid variant. PCR amplification results in a population of product DNA fragments that differ from the polynucleotide template encoding the polypeptide at the position specified by the primer. The product DNA fragments replace the corresponding region in the plasmid and this gives a polynucleotide encoding the desired amino acid variant.

A further technique for generating amino acid variants is the cassette mutagenesis technique described in Wells et al., *Gene* 34:315 (1985); and other mutagenesis techniques well known in the art, such as, for example, the techniques in Sambrook et al., supra, and *Current' Protocols in Molecular Biology*, Ausubel et al. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be used in the practice of the invention for the cloning and expression

of these novel nucleic acids. Such DNA sequences include those which are capable of hybridizing to the appropriate novel nucleic acid sequence under stringent conditions.

5

10

15

20

25

30

35

Polynucleotides encoding preferred polypeptide truncations of the invention can be used to generate polynucleotides encoding chimeric or fusion proteins comprising one or more domains of the invention and heterologous protein sequences.

The polynucleotides of the invention additionally include the complement of any of the polynucleotides recited above. The polynucleotide can be DNA (genomic, cDNA, amplified, or synthetic) or RNA. Methods and algorithms for obtaining such polynucleotides are well known to those of skill in the art and can include, for example, methods for determining hybridization conditions that can routinely isolate polynucleotides of the desired sequence identities.

In accordance with the invention, polynucleotide sequences comprising the mature protein coding sequences corresponding to any one of SEQ ID NO:1-1350, or functional equivalents thereof, may be used to generate recombinant DNA molecules that direct the expression of that nucleic acid, or a functional equivalent thereof, in appropriate host cells. Also included are the cDNA inserts of any of the clones identified herein.

A polynucleotide according to the invention can be joined to any of a variety of other nucleotide sequences by well-established recombinant DNA techniques (see Sambrook J et al. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY). Useful nucleotide sequences for joining to polynucleotides include an assortment of vectors, e.g., plasmids, cosmids, lambda phage derivatives, phagemids, and the like, that are well known in the art. Accordingly, the invention also provides a vector including a polynucleotide of the invention and a host cell containing the polynucleotide. In general, the vector contains an origin of replication functional in at least one organism, convenient restriction endonuclease sites, and a selectable marker for the host cell. Vectors according to the invention include expression vectors, replication vectors, probe generation vectors, and sequencing vectors. A host cell according to the invention can be a prokaryotic or eukaryotic cell and can be a unicellular organism or part of a multicellular organism.

The present invention further provides recombinant constructs comprising a nucleic acid having any of the nucleotide sequences of SEQ ID NO:1-1350 or a fragment thereof or any other polynucleotides of the invention. In one embodiment, the recombinant constructs of the present invention comprise a vector, such as a plasmid or viral vector, into which a nucleic acid having any of the nucleotide sequences of SEQ ID NO:1-1350 or a fragment thereof is inserted, in a forward or reverse orientation. In the case of a vector comprising one of the ORFs of the present invention, the vector may further comprise regulatory sequences, including for example, a promoter, operably linked to the ORF. Large numbers of suitable vectors and promoters are

known to those of skill in the art and are commercially available for generating the recombinant constructs of the present invention. The following vectors are provided by way of example. Bacterial: pBs, phagescript, PsiX174, pBluescript SK, pBs KS, pNH8a, pNH16a, pNH18a, pNH46a (Stratagene); pTrc99A, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia). Eukaryotic: pWLneo, pSV2cat, pOG44, PXTI, pSG (Stratagene) pSVK3, pBPV, pMSG, pSVL (Pharmacia).

The isolated polynucleotide of the invention may be operably linked to an expression control sequence such as the pMT2 or pED expression vectors disclosed in Kaufman et al., Nucleic Acids Res. 19, 4485-4490 (1991), in order to produce the protein recombinantly. Many suitable expression control sequences are known in the art. General methods of expressing recombinant proteins are also known and are exemplified in R. Kaufman, Methods in Enzymology 185, 537-566 (1990). As defined herein "operably linked" means that the isolated polynucleotide of the invention and an expression control sequence are situated within a vector or cell in such a way that the protein is expressed by a host cell which has been transformed (transfected) with the ligated polynucleotide/expression control sequence.

Promoter regions can be selected from any desired gene using CAT (chloramphenicol transferase) vectors or other vectors with selectable markers. Two appropriate vectors are pKK232-8 and pCM7. Particular named bacterial promoters include lacI, lacZ, T3, T7, gpt, lambda PR, and trc. Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein-I. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art.

Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the host cell, e.g., the ampicillin resistance gene of E. coli and S. cerevisiae TRP1 gene, and a promoter derived from a highly-expressed gene to direct transcription of a downstream structural sequence. Such promoters can be derived from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK), a-factor, acid phosphatase, or heat shock proteins, among others. The heterologous structural sequence is assembled in appropriate phase with translation initiation and termination sequences, and preferably, a leader sequence capable of directing secretion of translated protein into the periplasmic space or extracellular medium. Optionally, the heterologous sequence can encode a fusion protein including an amino terminal identification peptide imparting desired characteristics, e.g., stabilization or simplified purification of expressed recombinant product. Useful expression vectors for bacterial use are constructed by inserting a structural DNA sequence encoding a desired protein together with suitable translation initiation and termination signals in operable reading phase with a functional promoter. The vector will comprise one or

more phenotypic selectable markers and an origin of replication to ensure maintenance of the vector and to, if desirable, provide amplification within the host. Suitable prokaryotic hosts for transformation include *E. coli*, *Bacillus subtilis*, *Salmonella typhimurium* and various species within the genera *Pseudomonas*, *Streptomyces*, and *Staphylococcus*, although others may also be employed as a matter of choice.

As a representative but non-limiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017). Such commercial vectors include, for example, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and GEM 1 (Promega Biotech, Madison, WI, USA). These pBR322 "backbone" sections are combined with an appropriate promoter and the structural sequence to be expressed. Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter is induced or derepressed by appropriate means (e.g., temperature shift or chemical induction) and cells are cultured for an additional period. Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification.

Polynucleotides of the invention can also be used to induce immune responses. For example, as described in Fan et al., *Nat. Biotech.* 17:870-872 (1999), incorporated herein by reference, nucleic acid sequences encoding a polypeptide may be used to generate antibodies against the encoded polypeptide following topical administration of naked plasmid DNA or following injection, and preferably intramuscular injection of the DNA. The nucleic acid sequences are preferably inserted in a recombinant expression vector and may be in the form of naked DNA.

25 4.3 ANTISENSE

5

10

15

20

30

Another aspect of the invention pertains to isolated antisense nucleic acid molecules that are hybridizable to or complementary to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1-1350, or fragments, analogs or derivatives thereof. An "antisense" nucleic acid comprises a nucleotide sequence that is complementary to a "sense" nucleic acid encoding a protein, e.g., complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. In specific aspects, antisense nucleic acid molecules are provided that comprise a sequence complementary to at least about 10, 25, 50, 100, 250 or 500 nucleotides or an entire coding strand, or to only a portion thereof. Nucleic acid molecules encoding fragments, homologs, derivatives and analogs of a protein of any of SEO ID

NO:1351-2700 or antisense nucleic acids complementary to a nucleic acid sequence of SEQ ID NO:1-1350 are additionally provided.

In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence of the invention. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues. In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence of the invention. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (*i.e.*, also referred to as 5' and 3' untranslated regions).

5

10

15

20

Given the coding strand sequences encoding a nucleic acid disclosed herein (e.g., SEQ ID NO:1-1350), antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick or Hoogsteen base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of a mRNA, but more preferably is an oligonucleotide that is antisense to only a portion of the coding or noncoding region of a mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of a mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis or enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (e.g., an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, e.g., phosphorothioate derivatives and acridine substituted nucleotides can be used.

Examples of modified nucleotides that can be used to generate the antisense nucleic acid include: 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the

antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (*i.e.*, RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

5

10

15

20

25

30

35

The antisense nucleic acid molecules of the invention are typically administered to a subject or generated in situ such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a protein according to the invention to thereby inhibit expression of the protein, e.g., by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule that binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention includes direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, e.g., by linking the antisense nucleic acid molecules to peptides or antibodies that bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an -a nomeric nucleic acid molecule. An -a nomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual -units, the strands run parallel to each other (Gaultier et al. (1987) Nucleic Acids Res 15: 6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue et al. (1987) Nucleic Acids Res 15: 6131-6148) or a chimeric RNA -DNA analogue (Inoue et al. (1987) FEBS Lett 215: 327-330).

4.4 RIBOZYMES AND PNA MOIETIES

In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity that are capable of cleaving a single-stranded nucleic acid, such as a mRNA, to which they have a complementary region. Thus, ribozymes (e.g., hammerhead ribozymes (described in Haselhoff and Gerlach (1988) Nature 334:585-591)) can be used to catalytically cleave a mRNA transcripts to thereby inhibit translation of a mRNA. A ribozyme having specificity for a nucleic acid of the invention can be

designed based upon the nucleotide sequence of a DNA disclosed herein (i.e., SEQ ID NO:1-1350). For example, a derivative of a Tetrahymena L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a SECX-encoding mRNA. See, e.g., Cech et al. U.S. Pat. No. 4,987,071; and Cech et al. U.S. Pat. No. 5,116,742. Alternatively, SECX mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, e.g., Bartel et al., (1993) Science 261:1411-1418.

Alternatively, gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region (e.g., promoter and/or enhancers) to form triple helical structures that prevent transcription of the gene in target cells. See generally, Helene. (1991) Anticancer Drug Des. 6: 569-84; Helene. et al. (1992) Ann. N.Y. Acad. Sci. 660:27-36; and Maher (1992) Bioassays 14: 807-15.

10

. 15

20

25

30

35

In various embodiments, the nucleic acids of the invention can be modified at the base moiety, sugar moiety or phosphate backbone to improve, e.g., the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acids can be modified to generate peptide nucleic acids (see Hyrup et al. (1996) Bioorg Med Chem 4: 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, e.g., DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup et al. (1996) above; Perry-O'Keefe et al. (1996) PNAS 93: 14670-675.

PNAs of the invention can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, e.g., inducing transcription or translation arrest or inhibiting replication. PNAs of the invention can also be used, e.g., in the analysis of single base pair mutations in a gene by, e.g., PNA directed PCR clamping; as artificial restriction enzymes when used in combination with other enzymes, e.g., S1 nucleases (Hyrup B. (1996) above); or as probes or primers for DNA sequence and hybridization (Hyrup et al. (1996), above; Perry-O'Keefe (1996), above).

In another embodiment, PNAs of the invention can be modified, e.g., to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras can be generated that may

combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, e.g., RNase H and DNA polymerases, to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup (1996) above). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup (1996) above and Finn et al. (1996) Nucl Acids Res 24: 3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry, and modified nucleoside analogs, e.g., 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used between the PNA and the 5' end of DNA (Mag et al. (1989) Nucl Acid Res 17: 5973-88). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn et al. (1996) above). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment. See, Petersen et al. (1975) Bioorg Med Chem Lett 5: 1119-11124.

In other embodiments, the oligonucleotide may include other appended groups such as peptides (e.g., for targeting host cell receptors in vivo), or agents facilitating transport across the cell membrane (see, e.g., Letsinger et al., 1989, Proc. Natl. Acad. Sci. U.S.A. 86:6553-6556; Lemaitre et al., 1987, Proc. Natl. Acad. Sci. 84:648-652; PCT Publication No. W088/09810) or the blood-brain barrier (see, e.g., PCT Publication No. W089/10134). In addition, oligonucleotides can be modified with hybridization triggered cleavage agents (See, e.g., Krol et al., 1988, BioTechniques 6:958-976) or intercalating agents. (See, e.g., Zon, 1988, Pharm. Res. 5: 539-549). To this end, the oligonucleotide may be conjugated to another molecule, e.g., a peptide, a hybridization triggered cross-linking agent, a transport agent, a hybridization-triggered cleavage agent, etc.

25

30

35

10

15

20

4.5 HOSTS

The present invention further provides host cells genetically engineered to contain the polynucleotides of the invention. For example, such host cells may contain nucleic acids of the invention introduced into the host cell using known transformation, transfection or infection methods. The present invention still further provides host cells genetically engineered to express the polynucleotides of the invention, wherein such polynucleotides are in operative association with a regulatory sequence heterologous to the host cell which drives expression of the polynucleotides in the cell.

Knowledge of nucleic acid sequences allows for modification of cells to permit, or increase, expression of endogenous polypeptide. Cells can be modified (e.g., by homologous

recombination) to provide increased polypeptide expression by replacing, in whole or in part, the naturally occurring promoter with all or part of a heterologous promoter so that the cells express the polypeptide at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to the encoding sequences. See, for example, PCT International Publication No. WO94/12650, PCT International Publication No. WO92/20808, and PCT International Publication No. WO91/09955. It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (e.g., ada, dhfr, and the multifunctional CAD gene which encodes carbamyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the coding sequence, amplification of the marker DNA by standard selection methods results in coamplification of the desired protein coding sequences in the cells.

10

15

20

25

30

35

The host cell can be a higher eukaryotic host cell, such as a mammalian cell, a lower eukaryotic host cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. Introduction of the recombinant construct into the host cell can be effected by calcium phosphate transfection, DEAE, dextran mediated transfection, or electroporation (Davis, L et al., *Basic Methods in Molecular Biology* (1986)). The host cells containing one of the polynucleotides of the invention, can be used in conventional manners to produce the gene product encoded by the isolated fragment (in the case of an ORF) or can be used to produce a heterologous protein under the control of the EMF.

Any host/vector system can be used to express one or more of the ORFs of the present invention. These include, but are not limited to, eukaryotic hosts such as HeLa cells, Cv-1 cell, COS cells, 293 cells, and Sf9 cells, as well as prokaryotic host such as *E. coli* and *B. subtilis*. The most preferred cells are those which do not normally express the particular polypeptide or protein or which expresses the polypeptide or protein at low natural level. Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook, et al., in Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, New York (1989), the disclosure of which is hereby incorporated by reference.

Various mammalian cell culture systems can also be employed to express recombinant protein. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts, described by Gluzman, Cell 23:175 (1981). Other cell lines capable of expressing a compatible vector are, for example, the C127, monkey COS cells, Chinese Hamster Ovary (CHO) cells, human kidney 293 cells, human epidermal A431 cells, human Colo205 cells, 3T3

cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from *in vitro* culture of primary tissue, primary explants, HeLa cells, mouse L cells, BHK, HL-60, U937, HaK or Jurkat cells. Mammalian expression vectors will comprise an origin of replication, a suitable promoter and also any necessary ribosome binding sites, polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking nontranscribed sequences. DNA sequences derived from the SV40 viral genome, for example, SV40 origin, early promoter, enhancer, splice, and polyadenylation sites may be used to provide the required nontranscribed genetic elements. Recombinant polypeptides and proteins produced in bacterial culture are usually isolated by initial extraction from cell pellets, followed by one or more salting-out, aqueous ion exchange or size exclusion chromatography steps. Protein refolding steps can be used, as necessary, in completing configuration of the mature protein. Finally, high performance liquid chromatography (HPLC) can be employed for final purification steps. Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents.

Alternatively, it may be possible to produce the protein in lower eukaryotes such as yeast or insects or in prokaryotes such as bacteria. Potentially suitable yeast strains include Saccharomyces cerevisiae, Schizosaccharomyces pombe, Kluyveromyces strains, Candida, or any yeast strain capable of expressing heterologous proteins. Potentially suitable bacterial strains include Escherichia coli, Bacillus subtilis, Salmonella typhimurium, or any bacterial strain capable of expressing heterologous proteins. If the protein is made in yeast or bacteria, it may be necessary to modify the protein produced therein, for example by phosphorylation or glycosylation of the appropriate sites, in order to obtain the functional protein. Such covalent attachments may be accomplished using known chemical or enzymatic methods.

In another embodiment of the present invention, cells and tissues may be engineered to express an endogenous gene comprising the polynucleotides of the invention under the control of inducible regulatory elements, in which case the regulatory sequences of the endogenous gene may be replaced by homologous recombination. As described herein, gene targeting can be used to replace a gene's existing regulatory region with a regulatory sequence isolated from a different gene or a novel regulatory sequence synthesized by genetic engineering methods. Such regulatory sequences may be comprised of promoters, enhancers, scaffold-attachment regions, negative regulatory elements, transcriptional initiation sites, regulatory protein binding sites or combinations of said sequences. Alternatively, sequences which affect the structure or stability of the RNA or protein produced may be replaced, removed, added, or otherwise modified by targeting. These sequence include polyadenylation signals, mRNA stability elements, splice

sites, leader sequences for enhancing or modifying transport or secretion properties of the protein, or other sequences which alter or improve the function or stability of protein or RNA molecules.

5

10

15

20

25

30

35

The targeting event may be a simple insertion of the regulatory sequence, placing the gene under the control of the new regulatory sequence, e.g., inserting a new promoter or enhancer or both upstream of a gene. Alternatively, the targeting event may be a simple deletion of a regulatory element, such as the deletion of a tissue-specific negative regulatory element. Alternatively, the targeting event may replace an existing element; for example, a tissue-specific enhancer can be replaced by an enhancer that has broader or different cell-type specificity than the naturally occurring elements. Here, the naturally occurring sequences are deleted and new sequences are added. In all cases, the identification of the targeting event may be facilitated by the use of one or more selectable marker genes that are contiguous with the targeting DNA, allowing for the selection of cells in which the exogenous DNA has integrated into the host cell genome. The identification of the targeting event may also be facilitated by the use of one or more marker genes exhibiting the property of negative selection, such that the negatively selectable marker is linked to the exogenous DNA, but configured such that the negatively selectable marker flanks the targeting sequence, and such that a correct homologous recombination event with sequences in the host cell genome does not result in the stable integration of the negatively selectable marker. Markers useful for this purpose include the Herpes Simplex Virus thymidine kinase (TK) gene or the bacterial xanthine-guanine phosphoribosyl-transferase (gpt) gene.

The gene targeting or gene activation techniques which can be used in accordance with this aspect of the invention are more particularly described in U.S. Patent No. 5,272,071 to Chappel; U.S. Patent No. 5,578,461 to Sherwin et al.; International Application No. PCT/US92/09627 (WO93/09222) by Selden et al.; and International Application No. PCT/US90/06436 (WO91/06667) by Skoultchi et al., each of which is incorporated by reference herein in its entirety.

4.6 POLYPEPTIDES OF THE INVENTION

The isolated polypeptides of the invention include, but are not limited to, a polypeptide comprising: the amino acid sequences set forth as any one of SEQ ID NO:1351-2700 or an amino acid sequence encoded by any one of the nucleotide sequences SEQ ID NO:1-1350 or the corresponding full length or mature protein. Polypeptides of the invention also include polypeptides preferably with biological or immunological activity that are encoded by: (a) a polynucleotide having any one of the nucleotide sequences set forth in SEQ ID NO:1-1350 or (b)

polynucleotides encoding any one of the amino acid sequences set forth as SEQ ID NO:1351-2700 or (c) polynucleotides that hybridize to the complement of the polynucleotides of either (a) or (b) under stringent hybridization conditions. The invention also provides biologically active or immunologically active variants of any of the amino acid sequences set forth as SEQ ID NO:1351-2700 or the corresponding full length or mature protein; and "substantial equivalents" thereof (e.g., with at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, 86%, 87%, 88%, 89%, at least about 90%, 91%, 92%, 93%, 94%, typically at least about 95%, 96%, 97%, more typically at least about 98%, or most typically at least about 99% amino acid identity) that retain biological activity. Polypeptides encoded by allelic variants may have a similar, increased, or decreased activity compared to polypeptides comprising SEQ ID NO:1351-2700.

5

10

15

20

25

30

35

Fragments of the proteins of the present invention which are capable of exhibiting biological activity are also encompassed by the present invention. Fragments of the protein may be in linear form or they may be cyclized using known methods, for example, as described in H. U. Saragovi, et al., Bio/Technology 10, 773-778 (1992) and in R. S. McDowell, et al., J. Amer. Chem. Soc. 114, 9245-9253 (1992), both of which are incorporated herein by reference. Such fragments may be fused to carrier molecules such as immunoglobulins for many purposes, including increasing the valency of protein binding sites.

The present invention also provides both full-length and mature forms (for example, without a signal sequence or precursor sequence) of the disclosed proteins. The protein coding sequence is identified in the sequence listing by translation of the disclosed nucleotide sequences. The mature form of such protein may be obtained by expression of a full-length polynucleotide in a suitable mammalian cell or other host cell. The sequence of the mature form of the protein is also determinable from the amino acid sequence of the full-length form. Where proteins of the present invention are membrane bound, soluble forms of the proteins are also provided. In such forms, part or all of the regions causing the proteins to be membrane bound are deleted so that the proteins are fully secreted from the cell in which they are expressed.

Protein compositions of the present invention may further comprise an acceptable carrier, such as a hydrophilic, e.g., pharmaceutically acceptable, carrier.

The present invention further provides isolated polypeptides encoded by the nucleic acid fragments of the present invention or by degenerate variants of the nucleic acid fragments of the present invention. By "degenerate variant" is intended nucleotide fragments which differ from a nucleic acid fragment of the present invention (e.g., an ORF) by nucleotide sequence but, due to the degeneracy of the genetic code, encode an identical polypeptide sequence. Preferred nucleic acid fragments of the present invention are the ORFs that encode proteins.

A variety of methodologies known in the art can be utilized to obtain any one of the isolated polypeptides or proteins of the present invention. At the simplest level, the amino acid sequence can be synthesized using commercially available peptide synthesizers. The synthetically-constructed protein sequences, by virtue of sharing primary, secondary or tertiary structural and/or conformational characteristics with proteins may possess biological properties in common therewith, including protein activity. This technique is particularly useful in producing small peptides and fragments of larger polypeptides. Fragments are useful, for example, in generating antibodies against the native polypeptide. Thus, they may be employed as biologically active or immunological substitutes for natural, purified proteins in screening of therapeutic compounds and in immunological processes for the development of antibodies.

The polypeptides and proteins of the present invention can alternatively be purified from cells which have been altered to express the desired polypeptide or protein. As used herein, a cell is said to be altered to express a desired polypeptide or protein when the cell, through genetic manipulation, is made to produce a polypeptide or protein which it normally does not produce or which the cell normally produces at a lower level. One skilled in the art can readily adapt procedures for introducing and expressing either recombinant or synthetic sequences into eukaryotic or prokaryotic cells in order to generate a cell which produces one of the polypeptides or proteins of the present invention.

The invention also relates to methods for producing a polypeptide comprising growing a culture of host cells of the invention in a suitable culture medium, and purifying the protein from the cells or the culture in which the cells are grown. For example, the methods of the invention include a process for producing a polypeptide in which a host cell containing a suitable expression vector that includes a polynucleotide of the invention is cultured under conditions that allow expression of the encoded polypeptide. The polypeptide can be recovered from the culture, conveniently from the culture medium, or from a lysate prepared from the host cells and further purified. Preferred embodiments include those in which the protein produced by such process is a full length or mature form of the protein.

In an alternative method, the polypeptide or protein is purified from bacterial cells which naturally produce the polypeptide or protein. One skilled in the art can readily follow known methods for isolating polypeptides and proteins in order to obtain one of the isolated polypeptides or proteins of the present invention. These include, but are not limited to, immunochromatography, HPLC, size-exclusion chromatography, ion-exchange chromatography, and immuno-affinity chromatography. See, e.g., Scopes, Protein Purification: Principles and Practice, Springer-Verlag (1994); Sambrook, et al., in Molecular Cloning: A Laboratory Manual; Ausubel et al., Current Protocols in Molecular Biology. Polypeptide fragments that

retain biological/immunological activity include fragments comprising greater than about 100 amino acids, or greater than about 200 amino acids, and fragments that encode specific protein domains.

5

10

15

20

25

30

35

The purified polypeptides can be used in *in vitro* binding assays which are well known in the art to identify molecules which bind to the polypeptides. These molecules include but are not limited to, for *e.g.*, small molecules, molecules from combinatorial libraries, antibodies or other proteins. The molecules identified in the binding assay are then tested for antagonist or agonist activity in *in vivo* tissue culture or animal models that are well known in the art. In brief, the molecules are titrated into a plurality of cell cultures or animals and then tested for either cell/animal death or prolonged survival of the animal/cells.

In addition, the peptides of the invention or molecules capable of binding to the peptides may be complexed with toxins, e.g., ricin or cholera, or with other compounds that are toxic to cells. The toxin-binding molecule complex is then targeted to a tumor or other cell by the specificity of the binding molecule for SEQ ID NO:1351-2700.

The protein of the invention may also be expressed as a product of transgenic animals, e.g., as a component of the milk of transgenic cows, goats, pigs, or sheep which are characterized by somatic or germ cells containing a nucleotide sequence encoding the protein.

The proteins provided herein also include proteins characterized by amino acid sequences similar to those of purified proteins but into which modification are naturally provided or deliberately engineered. For example, modifications, in the peptide or DNA sequence, can be made by those skilled in the art using known techniques. Modifications of interest in the protein sequences may include the alteration, substitution, replacement, insertion or deletion of a selected amino acid residue in the coding sequence. For example, one or more of the cysteine residues may be deleted or replaced with another amino acid to alter the conformation of the molecule. Techniques for such alteration, substitution, replacement, insertion or deletion are well known to those skilled in the art (see, e.g., U.S. Pat. No. 4,518,584). Preferably, such alteration, substitution, replacement, insertion or deletion retains the desired activity of the protein. Regions of the protein that are important for the protein function can be determined by various methods known in the art including the alanine-scanning method which involved systematic substitution of single or strings of amino acids with alanine, followed by testing the resulting alanine-containing variant for biological activity. This type of analysis determines the importance of the substituted amino acid(s) in biological activity. Regions of the protein that are important for protein function may be determined by the eMATRIX program.

Other fragments and derivatives of the sequences of proteins which would be expected to retain protein activity in whole or in part and are useful for screening or other immunological

methodologies may also be easily made by those skilled in the art given the disclosures herein. Such modifications are encompassed by the present invention.

The protein may also be produced by operably linking the isolated polynucleotide of the invention to suitable control sequences in one or more insect expression vectors, and employing an insect expression system. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, e.g., Invitrogen, San Diego, Calif., U.S.A. (the MaxBatTM kit), and such methods are well known in the art, as described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987), incorporated herein by reference. As used herein, an insect cell capable of expressing a polynucleotide of the present invention is "transformed."

5

10

15

20

30

The protein of the invention may be prepared by culturing transformed host cells under culture conditions suitable to express the recombinant protein. The resulting expressed protein may then be purified from such culture (i.e., from culture medium or cell extracts) using known purification processes, such as gel filtration and ion exchange chromatography. The purification of the protein may also include an affinity column containing agents which will bind to the protein; one or more column steps over such affinity resins as concanavalin A-agarose, heparin-toyopearlTM or Cibacrom blue 3GA SepharoseTM; one or more steps involving hydrophobic interaction chromatography using such resins as phenyl ether, butyl ether, or propyl ether; or immunoaffinity chromatography.

Alternatively, the protein of the invention may also be expressed in a form which will facilitate purification. For example, it may be expressed as a fusion protein, such as those of maltose binding protein (MBP), glutathione-S-transferase (GST) or thioredoxin (TRX), or as a His tag. Kits for expression and purification of such fusion proteins are commercially available from New England BioLab (Beverly, Mass.), Pharmacia (Piscataway, N.J.) and Invitrogen, respectively. The protein can also be tagged with an epitope and subsequently purified by using a specific antibody directed to such epitope. One such epitope ("FLAG®") is commercially available from Kodak (New Haven, Conn.).

Finally, one or more reverse-phase high performance liquid chromatography (RP- HPLC) steps employing hydrophobic RP-HPLC media, e.g., silica gel having pendant methyl or other aliphatic groups, can be employed to further purify the protein. Some or all of the foregoing purification steps, in various combinations, can also be employed to provide a substantially homogeneous isolated recombinant protein. The protein thus purified is substantially free of other mammalian proteins and is defined in accordance with the present invention as an "isolated protein."

The polypeptides of the invention include analogs (variants). This embraces fragments, as well as peptides in which one or more amino acids has been deleted, inserted, or substituted. Also, analogs of the polypeptides of the invention embrace fusions of the polypeptides or modifications of the polypeptides of the invention, wherein the polypeptide or analog is fused to another moiety or moieties, *e.g.*, targeting moiety or another therapeutic agent. Such analogs may exhibit improved properties such as activity and/or stability. Examples of moieties which may be fused to the polypeptide or an analog include, for example, targeting moieties which provide for the delivery of polypeptide to pancreatic cells, *e.g.*, antibodies to pancreatic cells, antibodies to immune cells such as T-cells, monocytes, dendritic cells, granulocytes, etc., as well as receptor and ligands expressed on pancreatic or immune cells. Other moieties which may be fused to the polypeptide include therapeutic agents which are used for treatment, for example, immunosuppressive drugs such as cyclosporin, SK506, azathioprine, CD3 antibodies and steroids. Also, polypeptides may be fused to immune modulators, and other cytokines such as alpha or beta interferon.

15

35

10

5

4.6.1 DETERMINING POLYPEPTIDE AND POLYNUCLEOTIDE IDENTITY AND SIMILARITY

Preferred identity and/or similarity are designed to give the largest match between the sequences tested. Methods to determine identity and similarity are codified in computer 20 programs including, but are not limited to, the GCG program package, including GAP (Devereux, J., et al., Nucleic Acids Research 12(1):387 (1984); Genetics Computer Group, University of Wisconsin, Madison, WI), BLASTP, BLASTN, BLASTX, FASTA (Altschul, S.F. et al., J. Molec. Biol. 215:403-410 (1990), PSI-BLAST (Altschul S.F. et al., Nucleic Acids Res. vol. 25, pp. 3389-3402, herein incorporated by reference), eMatrix software (Wu et al., J. Comp. 25 Biol., Vol. 6, pp. 219-235 (1999), herein incorporated by reference), eMotif software (Nevill-Manning et al, ISMB-97, Vol. 4, pp. 202-209, herein incorporated by reference), pFam software (Sonnhammer et al., Nucleic Acids Res., Vol. 26(1), pp. 320-322 (1998), herein incorporated by reference) and the Kyte-Doolittle hydrophobocity prediction algorithm (J. Mol Biol, 157, pp. 105-31 (1982), incorporated herein by reference). The BLAST programs are publicly available 30 from the National Center for Biotechnology Information (NCBI) and other sources (BLAST Manual, Altschul, S., et al. NCB NLM NIH Bethesda, MD 20894; Altschul, S., et al., J. Mol. Biol. 215:403-410 (1990).

4.7 CHIMERIC AND FUSION PROTEINS

The invention also provides chimeric or fusion proteins. As used herein, a "chimeric protein" or "fusion protein" comprises a polypeptide of the invention operatively linked to

another polypeptide. Within a fusion protein the polypeptide according to the invention can correspond to all or a portion of a protein according to the invention. In one embodiment, a fusion protein comprises at least one biologically active portion of a protein according to the invention. In another embodiment, a fusion protein comprises at least two biologically active portions of a protein according to the invention. Within the fusion protein, the term "operatively linked" is intended to indicate that the polypeptide according to the invention and the other polypeptide are fused in-frame to each other. The polypeptide can be fused to the N-terminus or C-terminus.

5

10

15

20

25

30

35

For example, in one embodiment a fusion protein comprises a polypeptide according to the invention operably linked to the extracellular domain of a second protein.

In another embodiment, the fusion protein is a GST-fusion protein in which the polypeptide sequences of the invention are fused to the C-terminus of the GST (i.e., glutathione S-transferase) sequences.

In another embodiment, the fusion protein is an immunoglobulin fusion protein in which the polypeptide sequences according to the invention comprises one or more domains are fused to sequences derived from a member of the immunoglobulin protein family. The immunoglobulin fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject to inhibit an interaction between a ligand and a protein of the invention on the surface of a cell, to thereby suppress signal transduction *in vivo*. The immunoglobulin fusion proteins can be used to affect the bioavailability of a cognate ligand. Inhibition of the ligand/protein interaction may be useful therapeutically for both the treatment of proliferative and differentiative disorders, *e,g.*, cancer as well as modulating (*e.g.*, promoting or inhibiting) cell survival. Moreover, the immunoglobulin fusion proteins of the invention can be used as immunogens to produce antibodies in a subject, to purify ligands, and in screening assays to identify molecules that inhibit the interaction of a polypeptide of the invention with a ligand.

A chimeric or fusion protein of the invention can be produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, e.g., by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers that give rise to complementary overhangs between two consecutive gene fragments that can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for

example, Ausubel et al. (eds.) CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (e.g., a GST polypeptide). A nucleic acid encoding a polypeptide of the invention can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the protein of the invention.

4.8 GENE THERAPY

5

10

15

20

25

30

Mutations in the polynucleotides of the invention gene may result in loss of normal function of the encoded protein. The invention thus provides gene therapy to restore normal activity of the polypeptides of the invention; or to treat disease states involving polypeptides of the invention. Delivery of a functional gene encoding polypeptides of the invention to appropriate cells is effected ex vivo, in situ, or in vivo by use of vectors, and more particularly viral vectors (e.g., adenovirus, adeno-associated virus, or a retrovirus), or ex vivo by use of physical DNA transfer methods (e.g., liposomes or chemical treatments). See, for example, Anderson, Nature, supplement to vol. 392, no. 6679, pp.25-20 (1998). For additional reviews of gene therapy technology see Friedmann, Science, 244: 1275-1281 (1989); Verma, Scientific American: 68-84 (1990); and Miller, Nature, 357: 455-460 (1992). Introduction of any one of the nucleotides of the present invention or a gene encoding the polypeptides of the present invention can also be accomplished with extrachromosomal substrates (transient expression) or artificial chromosomes (stable expression). Cells may also be cultured ex vivo in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced in vivo for therapeutic purposes. Alternatively, it is contemplated that in other human disease states, preventing the expression of or inhibiting the activity of polypeptides of the invention will be useful in treating the disease states. It is contemplated that antisense therapy or gene therapy could be applied to negatively regulate the expression of polypeptides of the invention.

Other methods inhibiting expression of a protein include the introduction of antisense molecules to the nucleic acids of the present invention, their complements, or their translated RNA sequences, by methods known in the art. Further, the polypeptides of the present invention can be inhibited by using targeted deletion methods, or the insertion of a negative regulatory element such as a silencer, which is tissue specific.

The present invention still further provides cells genetically engineered *in vivo* to express the polynucleotides of the invention, wherein such polynucleotides are in operative association with a regulatory sequence heterologous to the host cell which drives expression of the polynucleotides in

the cell. These methods can be used to increase or decrease the expression of the polynucleotides of the present invention.

Knowledge of DNA sequences provided by the invention allows for modification of cells to permit, increase, or decrease, expression of endogenous polypeptide. Cells can be modified (e.g., by homologous recombination) to provide increased polypeptide expression by replacing, in whole or in part, the naturally occurring promoter with all or part of a heterologous promoter so that the cells express the protein at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to the desired protein encoding sequences. See, for example, PCT International Publication No. WO 94/12650, PCT International Publication No. WO 92/20808, and PCT International Publication No. WO 91/09955. It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (e.g., ada, dhfr, and the multifunctional CAD gene which encodes carbamyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the desired protein coding sequence, amplification of the marker DNA by standard selection methods results in co-amplification of the desired protein coding sequences in the cells.

In another embodiment of the present invention, cells and tissues may be engineered to express an endogenous gene comprising the polynucleotides of the invention under the control of inducible regulatory elements, in which case the regulatory sequences of the endogenous gene may be replaced by homologous recombination. As described herein, gene targeting can be used to replace a gene's existing regulatory region with a regulatory sequence isolated from a different gene or a novel regulatory sequence synthesized by genetic engineering methods. Such regulatory sequences may be comprised of promoters, enhancers, scaffold-attachment regions, negative regulatory elements, transcriptional initiation sites, regulatory protein binding sites or combinations of said sequences. Alternatively, sequences which affect the structure or stability of the RNA or protein produced may be replaced, removed, added, or otherwise modified by targeting. These sequences include polyadenylation signals, mRNA stability elements, splice sites, leader sequences for enhancing or modifying transport or secretion properties of the protein, or other sequences which after or improve the function or stability of protein or RNA molecules.

The targeting event may be a simple insertion of the regulatory sequence, placing the gene under the control of the new regulatory sequence, e.g., inserting a new promoter or enhancer or both upstream of a gene. Alternatively, the targeting event may be a simple deletion of a regulatory element, such as the deletion of a tissue-specific negative regulatory element. Alternatively, the targeting event may replace an existing element; for example, a tissue-specific enhancer can be replaced by an enhancer that has broader or different cell-type specificity than the naturally occurring elements. Here, the naturally occurring sequences are deleted and new sequences are

added. In all cases, the identification of the targeting event may be facilitated by the use of one or more selectable marker genes that are contiguous with the targeting DNA, allowing for the selection of cells in which the exogenous DNA has integrated into the cell genome. The identification of the targeting event may also be facilitated by the use of one or more marker genes exhibiting the property of negative selection, such that the negatively selectable marker is linked to the exogenous DNA, but configured such that the negatively selectable marker flanks the targeting sequence, and such that a correct homologous recombination event with sequences in the host cell genome does not result in the stable integration of the negatively selectable marker. Markers useful for this purpose include the Herpes Simplex Virus thymidine kinase (TK) gene or the bacterial xanthine-guanine phosphoribosyl-transferase (gpt) gene.

The gene targeting or gene activation techniques which can be used in accordance with this aspect of the invention are more particularly described in U.S. Patent No. 5,272,071 to Chappel; U.S. Patent No. 5,578,461 to Sherwin et al.; International Application No. PCT/US92/09627 (WO93/09222) by Selden et al.; and International Application No. PCT/US90/06436 (WO91/06667) by Skoultchi et al., each of which is incorporated by reference herein in its entirety.

4.9 TRANSGENIC ANIMALS

5

10

15

20

25

30

35

In preferred methods to determine biological functions of the polypeptides of the invention in vivo, one or more genes provided by the invention are either over expressed or inactivated in the germ line of animals using homologous recombination [Capecchi, Science 244:1288-1292 (1989)]. Animals in which the gene is over expressed, under the regulatory control of exogenous or endogenous promoter elements, are known as transgenic animals. Animals in which an endogenous gene has been inactivated by homologous recombination are referred to as "knockout" animals. Knockout animals, preferably non-human mammals, can be prepared as described in U.S. Patent No. 5,557,032, incorporated herein by reference. Transgenic animals are useful to determine the roles polypeptides of the invention play in biological processes, and preferably in disease states. Transgenic animals are useful as model systems to identify compounds that modulate lipid metabolism. Transgenic animals, preferably non-human mammals, are produced using methods as described in U.S. Patent No 5,489,743 and PCT Publication No. WO94/28122, incorporated herein by reference.

Transgenic animals can be prepared wherein all or part of a promoter of the polynucleotides of the invention is either activated or inactivated to alter the level of expression of the polypeptides of the invention. Inactivation can be carried out using homologous recombination methods described above. Activation can be achieved by supplementing or even replacing the homologous promoter to provide for increased protein expression. The homologous

promoter can be supplemented by insertion of one or more heterologous enhancer elements known to confer promoter activation in a particular tissue.

The polynucleotides of the present invention also make possible the development, through, e.g., homologous recombination or knock out strategies, of animals that fail to express polypeptides of the invention or that express a variant polypeptide. Such animals are useful as models for studying the *in vivo* activities of polypeptide as well as for studying modulators of the polypeptides of the invention.

5

10

15

20

25

30

35

In preferred methods to determine biological functions of the polypeptides of the invention *in vivo*, one or more genes provided by the invention are either over expressed or inactivated in the germ line of animals using homologous recombination [Capecchi, Science 244:1288-1292 (1989)]. Animals in which the gene is over expressed, under the regulatory control of exogenous or endogenous promoter elements, are known as transgenic animals. Animals in which an endogenous gene has been inactivated by homologous recombination are referred to as "knockout" animals. Knockout animals, preferably non-human mammals, can be prepared as described in U.S. Patent No. 5,557,032, incorporated herein by reference. Transgenic animals are useful to determine the roles polypeptides of the invention play in biological processes, and preferably in disease states. Transgenic animals are useful as model systems to identify compounds that modulate lipid metabolism. Transgenic animals, preferably non-human mammals, are produced using methods as described in U.S. Patent No 5,489,743 and PCT Publication No. WO94/28122, incorporated herein by reference.

Transgenic animals can be prepared wherein all or part of the polynucleotides of the invention promoter is either activated or inactivated to alter the level of expression of the polypeptides of the invention. Inactivation can be carried out using homologous recombination methods described above. Activation can be achieved by supplementing or even replacing the homologous promoter to provide for increased protein expression. The homologous promoter can be supplemented by insertion of one or more heterologous enhancer elements known to confer promoter activation in a particular tissue.

4.10 USES AND BIOLOGICAL ACTIVITY

The polynucleotides and proteins of the present invention are expected to exhibit one or more of the uses or biological activities (including those associated with assays cited herein) identified herein. Uses or activities described for proteins of the present invention may be provided by administration or use of such proteins or of polynucleotides encoding such proteins (such as, for example, in gene therapies or vectors suitable for introduction of DNA). The mechanism underlying the particular condition or pathology will dictate whether the

polypeptides of the invention, the polynucleotides of the invention or modulators (activators or inhibitors) thereof would be beneficial to the subject in need of treatment. Thus, "therapeutic compositions of the invention" include compositions comprising isolated polynucleotides (including recombinant DNA molecules, cloned genes and degenerate variants thereof) or polypeptides of the invention (including full length protein, mature protein and truncations or domains thereof), or compounds and other substances that modulate the overall activity of the target gene products, either at the level of target gene/protein expression or target protein activity. Such modulators include polypeptides, analogs, (variants), including fragments and fusion proteins, antibodies and other binding proteins; chemical compounds that directly or indirectly activate or inhibit the polypeptides of the invention (identified, e.g., via drug screening assays as described herein); antisense polynucleotides and polynucleotides suitable for triple helix formation; and in particular antibodies or other binding partners that specifically recognize one or more epitopes of the polypeptides of the invention.

The polypeptides of the present invention may likewise be involved in cellular activation or in one of the other physiological pathways described herein.

4.10.1 RESEARCH USES AND UTILITIES

5

10

15

20

25

30

35

The polynucleotides provided by the present invention can be used by the research community for various purposes. The polynucleotides can be used to express recombinant protein for analysis, characterization or therapeutic use; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as molecular weight markers on gels; as chromosome markers or tags (when labeled) to identify chromosomes or to map related gene positions; to compare with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of discovering other novel polynucleotides; for selecting and making oligomers for attachment to a "gene chip" or other support, including for examination of expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the polynucleotide can also be used in interaction trap assays (such as, for example, that described in Gyuris et al., Cell 75:791-803 (1993)) to identify polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

The polypeptides provided by the present invention can similarly be used in assays to determine biological activity, including in a panel of multiple proteins for high-throughput screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine levels of the protein (or its receptor) in biological fluids; as markers for tissues in which the corresponding polypeptide is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state); and, of course, to isolate correlative receptors or ligands. Proteins involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these research utilities are capable of being developed into reagent grade or kit format for commercialization as research products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E. F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S. L. and A. R. Kimmel eds., 1987.

4.10.2 NUTRITIONAL USES

5

10

15

20

25

30

35

Polynucleotides and polypeptides of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein or amino acid supplement, use as a carbon source, use as a nitrogen source and use as a source of carbohydrate. In such cases the polypeptide or polynucleotide of the invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the polypeptide or polynucleotide of the invention can be added to the medium in or on which the microorganism is cultured.

4.10.3 CYTOKINE AND CELL PROLIFERATION/DIFFERENTIATION ACTIVITY

A polypeptide of the present invention may exhibit activity relating to cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Many protein factors discovered to date, including all known cytokines, have exhibited activity in one or more factor-dependent cell proliferation assays, and hence the assays serve as a convenient

confirmation of cytokine activity. The activity of therapeutic compositions of the present invention is evidenced by any one of a number of routine factor dependent cell proliferation assays for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+(preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e, CMK, HUVEC, and Caco. Therapeutic compositions of the invention can be used in the following:

Assays for T-cell or thymocyte proliferation include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, *In Vitro* assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Bertagnolli et al., J. Immunol. 145:1706-1712, 1990; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Bertagnolli, et al., I. Immunol. 149:3778-3783, 1992; Bowman et al., I. Immunol. 152:1756-1761, 1994.

10

35

Assays for cytokine production and/or proliferation of spleen cells, lymph node cells or thymocytes include, without limitation, those described in: Polyclonal T cell stimulation, Kruisbeek, A. M. and Shevach, E. M. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 3.12.1-3.12.14, John Wiley and Sons, Toronto. 1994; and Measurement of mouse and human interleukin-γ, Schreiber, R. D. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 6.8.1-6.8.8, John Wiley and Sons, Toronto. 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic cells 20 include, without limitation, those described in: Measurement of Human and Murine Interleukin 2 and Interleukin 4, Bottomly, K., Davis, L. S. and Lipsky, P. E. In Current Protocols in Immunology, J. E. e.a. Coligan eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto, 1991: deVries et al., J. Exp. Med. 173:1205-1211, 1991; Moreau et al., Nature 336:690-692, 1988; Greenberger et al., Proc. Natl. Acad. Sci. U.S.A. 80:2931-2938, 1983; Measurement of mouse 25 and human interleukin 6--Nordan, R. In Current Protocols in Immunology, J. E. Coligan eds. Vol 1 pp. 6.6.1-6.6.5, John Wiley and Sons, Toronto. 1991; Smith et al., Proc. Natl. Aced. Sci. U.S.A. 83:1857-1861, 1986; Measurement of human Interleukin 11--Bennett, F., Giannotti, J., Clark, S. C. and Turner, K. J. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto. 1991; Measurement of mouse and human Interleukin 9--Ciarletta, A., Giannotti, J., Clark, S. C. and Turner, K. J. In Current Protocols in Immunology. 30 J. E. Coligan eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto. 1991.

Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and cytokine production) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W Strober,

Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function; Chapter 6, Cytokines and their cellular receptors; Chapter 7, Immunologic studies in Humans); Weinberger et al., Proc. Natl. Acad. Sci. USA 77:6091-6095, 1980; Weinberger et al., Eur. J. Immun. 11:405-411, 1981; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988.

4.10.4 STEM CELL GROWTH FACTOR ACTIVITY

5

10

20

25

30

35

A polypeptide of the present invention may exhibit stem cell growth factor activity and be involved in the proliferation, differentiation and survival of pluripotent and totipotent stem cells including primordial germ cells, embryonic stem cells, hematopoietic stem cells and/or germ line stem cells. Administration of the polypeptide of the invention to stem cells in vivo or ex vivo is expected to maintain and expand cell populations in a totipotential or pluripotential state which would be useful for re-engineering damaged or diseased tissues, transplantation, manufacture of bio-pharmaceuticals and the development of bio-sensors. The ability to produce 15 large quantities of human cells has important working applications for the production of human proteins which currently must be obtained from non-human sources or donors, implantation of cells to treat diseases such as Parkinson's, Alzheimer's and other neurodegenerative diseases; tissues for grafting such as bone marrow, skin, cartilage, tendons, bone, muscle (including cardiac muscle), blood vessels, cornea, neural cells, gastrointestinal cells and others; and organs for transplantation such as kidney, liver, pancreas (including islet cells), heart and lung.

It is contemplated that multiple different exogenous growth factors and/or cytokines may be administered in combination with the polypeptide of the invention to achieve the desired effect, including any of the growth factors listed herein, other stem cell maintenance factors, and specifically including stem cell factor (SCF), leukemia inhibitory factor (LIF), Flt-3 ligand (Flt-3L), any of the interleukins, recombinant soluble IL-6 receptor fused to IL-6, macrophage inflammatory protein 1-alpha (MIP-1-alpha), G-CSF, GM-CSF, thrombopoietin (TPO), platelet factor 4 (PF-4), platelet-derived growth factor (PDGF), neural growth factors and basic fibroblast growth factor (bFGF).

Since totipotent stem cells can give rise to virtually any mature cell type, expansion of these cells in culture will facilitate the production of large quantities of mature cells. Techniques for culturing stem cells are known in the art and administration of polypeptides of the invention, optionally with other growth factors and/or cytokines, is expected to enhance the survival and proliferation of the stem cell populations. This can be accomplished by direct administration of the polypeptide of the invention to the culture medium. Alternatively, stroma cells transfected with a polynucleotide that encodes for the polypeptide of the invention can be used as a feeder

layer for the stem cell populations in culture or in vivo. Stromal support cells for feeder layers may include embryonic bone marrow fibroblasts, bone marrow stromal cells, fetal liver cells, or cultured embryonic fibroblasts (see U.S. Patent No. 5,690,926).

5

10

15

20

25

30

35

Stem cells themselves can be transfected with a polynucleotide of the invention to induce autocrine expression of the polypeptide of the invention. This will allow for generation of undifferentiated totipotential/pluripotential stem cell lines that are useful as is or that can then be differentiated into the desired mature cell types. These stable cell lines can also serve as a source of undifferentiated totipotential/pluripotential mRNA to create cDNA libraries and templates for polymerase chain reaction experiments. These studies would allow for the isolation and identification of differentially expressed genes in stem cell populations that regulate stem cell proliferation and/or maintenance.

Expansion and maintenance of totipotent stem cell populations will be useful in the treatment of many pathological conditions. For example, polypeptides of the present invention may be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be used to augment or replace cells damaged by illness, autoimmune disease, accidental damage or genetic disorders. The polypeptide of the invention may be useful for inducing the proliferation of neural cells and for the regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders which involve degeneration, death or trauma to neural cells or nerve tissue. In addition, the expanded stem cell populations can also be genetically altered for gene therapy purposes and to decrease host rejection of replacement tissues after grafting or implantation.

Expression of the polypeptide of the invention and its effect on stem cells can also be manipulated to achieve controlled differentiation of the stem cells into more differentiated cell types. A broadly applicable method of obtaining pure populations of a specific differentiated cell type from undifferentiated stem cell populations involves the use of a cell-type specific promoter driving a selectable marker. The selectable marker allows only cells of the desired type to survive. For example, stem cells can be induced to differentiate into cardiomyocytes (Wobus et al., Differentiation, 48: 173-182, (1991); Klug et al., J. Clin. Invest., 98(1): 216-224, (1998)) or skeletal muscle cells (Browder, L. W. In: *Principles of Tissue Engineering eds*. Lanza et al., Academic Press (1997)). Alternatively, directed differentiation of stem cells can be accomplished by culturing the stem cells in the presence of a differentiation factor such as retinoic acid and an antagonist of the polypeptide of the invention which would inhibit the effects of endogenous stem cell factor activity and allow differentiation to proceed.

In vitro cultures of stem cells can be used to determine if the polypeptide of the invention exhibits stem cell growth factor activity. Stem cells are isolated from any one of various cell

sources (including hematopoietic stem cells and embryonic stem cells) and cultured on a feeder layer, as described by Thompson et al. Proc. Natl. Acad. Sci, U.S.A., 92: 7844-7848 (1995), in the presence of the polypeptide of the invention alone or in combination with other growth factors or cytokines. The ability of the polypeptide of the invention to induce stem cells proliferation is determined by colony formation on semi-solid support e.g. as described by Bernstein et al., Blood, 77: 2316-2321 (1991).

4.10.5 HEMATOPOIESIS REGULATING ACTIVITY

5

10

15

20

25

30

35

A polypeptide of the present invention may be involved in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell disorders. Even marginal biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis, e.g. in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (i.e., traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either in-vivo or ex-vivo (i.e., in conjunction with bone marrow transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)). as normal cells or genetically manipulated for gene therapy.

Therapeutic compositions of the invention can be used in the following:

Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without limitation, those described in: Johansson et al. Cellular Biology 15:141-151, 1995; Keller et al., Molecular and Cellular Biology 13:473-486, 1993; McClanahan et al., Blood 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, M. G. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 265-268, Wiley-Liss, Inc., New York, N.Y. 1994; Hirayama et al., Proc. Natl. Acad. Sci. USA 89:5907-5911, 1992; Primitive hematopoietic colony forming cells with high proliferative potential, McNiece, I. K. and Briddell, R. A. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 23-39, Wiley-Liss, Inc., New York, N.Y. 1994; Neben et al., Experimental Hematology 22:353-359, 1994; Cobblestone area forming cell assay, Ploemacher, R. E. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 1-21, Wiley-Liss, Inc., New York, N.Y. 1994; Long term bone marrow cultures in the presence of stromal cells, Spooncer, E., Dexter, M. and Allen, T. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 163-179, Wiley-Liss, Inc., New York, N.Y. 1994; Long term culture initiating cell assay, Sutherland, H. J. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 139-162, Wiley-Liss, Inc., New York, N.Y. 1994.

15

20

25

30

10

4.10.6 TISSUE GROWTH ACTIVITY

A polypeptide of the present invention also may be involved in bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as in wound healing and tissue repair and replacement, and in healing of burns, incisions and ulcers.

A polypeptide of the present invention which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Compositions of a polypeptide, antibody, binding partner, or other modulator of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. De novo bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A polypeptide of this invention may also be involved in attracting bone-forming cells, stimulating growth of bone-forming cells, or inducing differentiation of progenitors of bone-forming cells. Treatment of osteoporosis, osteoarthritis, bone degenerative disorders, or periodontal disease, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes may also be possible using the composition of the invention.

5

10

15

20

25

30

35

Another category of tissue regeneration activity that may involve the polypeptide of the present invention is tendon/ligament formation. Induction of tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. De novo tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors ex vivo for return in vivo to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

The compositions of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a composition may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a composition of the invention.

Compositions of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

Compositions of the present invention may also be involved in the generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine,

kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring may allow normal tissue to regenerate. A polypeptide of the present invention may also exhibit angiogenic activity.

A composition of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A composition of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

Therapeutic compositions of the invention can be used in the following:

Assays for tissue generation activity include, without limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No. WO91/07491 (skin, endothelium).

Assays for wound healing activity include, without limitation, those described in: Winter, Epidermal Wound Healing, pps. 71-112 (Maibach, H. I. and Rovee, D. T., eds.), Year Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, J. Invest. Dermatol 71:382-84 (1978).

20

25

30

5

10

15

4.10.7 IMMUNE STIMULATING OR SUPPRESSING ACTIVITY

A polypeptide of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A polynucleotide of the invention can encode a polypeptide exhibiting such activities. A protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), e.g., in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (e.g., HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious diseases causes by viral, bacterial, fungal or other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis viruses, herpes viruses, mycobacteria, Leishmania spp., malaria spp. and various fungal infections such as candidiasis. Of course, in this regard, proteins of the present invention may also be useful where a boost to the immune system generally may be desirable, i.e., in the treatment of cancer.

5

10

15

20

25

30

Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitis, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein (or antagonists thereof, including antibodies) of the present invention may also to be useful in the treatment of allergic reactions and conditions (e.g., anaphylaxis, serum sickness, drug reactions, food allergies, insect venom allergies, mastocytosis, allergic rhinitis, hypersensitivity pneumonitis, urticaria, angioedema, eczema, atopic dermatitis, allergic contact dermatitis, erythema multiforme, Stevens-Johnson syndrome, allergic conjunctivitis, atopic keratoconjunctivitis, venereal keratoconjunctivitis, giant papillary conjunctivitis and contact allergies), such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for example, organ transplantation), may also be treatable using a protein (or antagonists thereof) of the present invention. The therapeutic effects of the polypeptides or antagonists thereof on allergic reactions can be evaluated by in vivo animals models such as the cumulative contact enhancement test (Lastbom et al., Toxicology 125: 59-66, 1998), skin prick test (Hoffmann et al., Allergy 54: 446-54, 1999), guinea pig skin sensitization test (Vohr et al., Arch. Toxocol. 73: 501-9), and murine local lymph node assay (Kimber et al., J. Toxicol. Environ. Health 53: 563-79).

Using the proteins of the invention it may also be possible to modulate immune responses, in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response. The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or anergy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as, for example, B7)), e.g., preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue

transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a therapeutic composition of the invention may prevent cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, a lack of costimulation may also be sufficient to anergize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or tolerance in a subject, it may also be necessary to block the function of a combination of B lymphocyte antigens.

The efficacy of particular therapeutic compositions in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the immunosuppressive effects of CTLA4Ig fusion proteins in vivo as described in Lenschow et al., Science 257:789-792 (1992) and Turka et al., Proc. Natl. Acad. Sci USA, 89:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of therapeutic compositions of the invention on the development of that disease.

10

15

20

25

30

35

Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and autoantibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block stimulation of T cells can be used to inhibit T cell activation and prevent production of autoantibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythmatosis in MRL/lpr/lpr mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (e.g., a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial

immune response. For example, enhancing an immune response may be useful in cases of viral infection, including systemic viral diseases such as influenza, the common cold, and encephalitis.

5

10

15

20

25

30

35

Alternatively, anti-viral immune responses may be enhanced in an infected patient by removing T cells from the patient, costimulating the T cells in vitro with viral antigen-pulsed APCs either expressing a peptide of the present invention or together with a stimulatory form of a soluble peptide of the present invention and reintroducing the in vitro activated T cells into the patient. Another method of enhancing anti-viral immune responses would be to isolate infected cells from a patient, transfect them with a nucleic acid encoding a protein of the present invention as described herein such that the cells express all or a portion of the protein on their surface, and reintroduce the transfected cells into the patient. The infected cells would now be capable of delivering a costimulatory signal to, and thereby activate, T cells in vivo.

A polypeptide of the present invention may provide the necessary stimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells. In addition, tumor cells which lack MHC class I or MHC class II molecules, or which fail to reexpress sufficient mounts of MHC class I or MHC class II molecules, can be transfected with nucleic acid encoding all or a portion of (e.g., a cytoplasmic-domain truncated portion) of an MHC class I alpha chain protein and β_2 microglobulin protein or an MHC class II alpha chain protein and an MHC class II beta chain protein to thereby express MHC class I or MHC class II proteins on the cell surface. Expression of the appropriate class I or class II MHC in conjunction

with a peptide having the activity of a B lymphocyte antigen (e.g., B7-1, B7-2, B7-3) induces a T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor specific immunity. Thus, the induction of a T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for thymocyte or splenocyte cytotoxicity include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J.

Immunol. 135:1564-1572, 1985; Takai et al., I. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bowman et al., J. Virology 61:1992-1998; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Brown et al., J. Immunol. 153:3079-3092, 1994.

Assays for T-cell-dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, J. Immunol. 144:3028-3033, 1990; and Assays for B cell function: In vitro antibody production, Mond, J. J. and Brunswick, M. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto. 1994.

5

10

15

20

25

30

35

Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., J. Immunol. 149:3778-3783, 1992.

Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery et al., J. Immunol. 134:536-544, 1995; Inaba et al., Journal of Experimental Medicine 173:549-559, 1991; Macatonia et al., Journal of Immunology 154:5071-5079, 1995; Porgador et al., Journal of Experimental Medicine 182:255-260, 1995; Nair et al., Journal of Virology 67:4062-4069, 1993; Huang et al., Science 264:961-965, 1994; Macatonia et al., Journal of Experimental Medicine 169:1255-1264, 1989; Bhardwaj et al., Journal of Clinical Investigation 94:797-807, 1994; and Inaba et al., Journal of Experimental Medicine 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz et al., Cytometry 13:795-808, 1992; Gorczyca et al., Leukemia 7:659-670, 1993; Gorczyca et al., Cancer Research 53:1945-1951, 1993; Itoh et al., Cell 66:233-243, 1991; Zacharchuk, Journal of Immunology 145:4037-4045, 1990; Zamai et al., Cytometry 14:891-897, 1993; Gorczyca et al., International Journal of Oncology 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica et al., Blood 84:111-117, 1994; Fine et al., Cellular Immunology 155:111-122, 1994; Galy et al., Blood 85:2770-2778, 1995; Toki et al., Proc. Nat. Acad Sci. USA 88:7548-7551, 1991.

4.10.8 ACTIVIN/INHIBIN ACTIVITY

5

10

15

20

30

35

A polypeptide of the present invention may also exhibit activin- or inhibin-related activities. A polynucleotide of the invention may encode a polypeptide exhibiting such characteristics. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins and are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a polypeptide of the present invention, alone or in heterodimers with a member of the inhibin family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the polypeptide of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, U.S. Pat. No. 4,798,885. A polypeptide of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as, but not limited to, cows, sheep and pigs.

The activity of a polypeptide of the invention may, among other means, be measured by the following methods.

Assays for activin/inhibin activity include, without limitation, those described in: Vale et al., Endocrinology 91:562-572, 1972; Ling et al., Nature 321:779-782, 1986; Vale et al., Nature 321:776-779, 1986; Mason et al., Nature 318:659-663, 1985; Forage et al., Proc. Natl. Acad. Sci. USA 83:3091-3095, 1986.

25 4.10.9 CHEMOTACTIC/CHEMOKINETIC ACTIVITY

A polypeptide of the present invention may be involved in chemotactic or chemokinetic activity for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Chemotactic and chemokinetic receptor activation can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic compositions (e.g. proteins, antibodies, binding partners, or modulators of the invention) provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

Therapeutic compositions of the invention can be used in the following:

5

10

15

20

25

30

35

Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Marguiles, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 6.12, Measurement of alpha and beta Chemokines 6.12.1-6.12.28; Taub et al. J. Clin. Invest. 95:1370-1376, 1995; Lind et al. APMIS 103:140-146, 1995; Muller et al Eur. J. Immunol. 25:1744-1748; Gruber et al. J. of Immunol. 152:5860-5867, 1994; Johnston et al. J. of Immunol. 153:1762-1768, 1994.

4.10.10 HEMOSTATIC AND THROMBOLYTIC ACTIVITY

A polypeptide of the invention may also be involved in hemostatis or thrombolysis or thrombosis. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Compositions may be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A composition of the invention may also be useful for dissolving or inhibiting formation of thromboses and for treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central nervous system vessels (e.g., stroke).

Therapeutic compositions of the invention can be used in the following:

Assay for hemostatic and thrombolytic activity include, without limitation, those described in: Linet et al., J. Clin. Pharmacol. 26:131-140, 1986; Burdick et al., Thrombosis Res. 45:413-419, 1987; Humphrey et al., Fibrinolysis 5:71-79 (1991); Schaub, Prostaglandins 35:467-474, 1988.

4.10.11 CANCER DIAGNOSIS AND THERAPY

Polypeptides of the invention may be involved in cancer cell generation, proliferation or metastasis. Detection of the presence or amount of polynucleotides or polypeptides of the

invention may be useful for the diagnosis and/or prognosis of one or more types of cancer. For example, the presence or increased expression of a polynucleotide/polypeptide of the invention may indicate a hereditary risk of cancer, a precancerous condition, or an ongoing malignancy. Conversely, a defect in the gene or absence of the polypeptide may be associated with a cancer condition. Identification of single nucleotide polymorphisms associated with cancer or a predisposition to cancer may also be useful for diagnosis or prognosis.

5

10

15

20

25

30

35

Cancer treatments promote tumor regression by inhibiting tumor cell proliferation, inhibiting angiogenesis (growth of new blood vessels that is necessary to support tumor growth) and/or prohibiting metastasis by reducing tumor cell motility or invasiveness. Therapeutic compositions of the invention may be effective in adult and pediatric oncology including in solid phase tumors/malignancies, locally advanced tumors, human soft tissue sarcomas, metastatic cancer, including lymphatic metastases, blood cell malignancies including multiple myeloma, acute and chronic leukemias, and lymphomas, head and neck cancers including mouth cancer, larynx cancer and thyroid cancer, lung cancers including small cell carcinoma and non-small cell cancers, breast cancers including small cell carcinoma and ductal carcinoma, gastrointestinal cancers including esophageal cancer, stomach cancer, colon cancer, colorectal cancer and polyps associated with colorectal neoplasia, pancreatic cancers, liver cancer, urologic cancers including bladder cancer and prostate cancer, malignancies of the female genital tract including ovarian carcinoma, uterine (including endometrial) cancers, and solid tumor in the ovarian follicle. kidney cancers including renal cell carcinoma, brain cancers including intrinsic brain tumors, neuroblastoma, astrocytic brain tumors, gliomas, metastatic tumor cell invasion in the central nervous system, bone cancers including osteomas, skin cancers including malignant melanoma, tumor progression of human skin keratinocytes, squamous cell carcinoma, basal cell carcinoma, hemangiopericytoma and Karposi's sarcoma.

Polypeptides, polynucleotides, or modulators of polypeptides of the invention (including inhibitors and stimulators of the biological activity of the polypeptide of the invention) may be administered to treat cancer. Therapeutic compositions can be administered in therapeutically effective dosages alone or in combination with adjuvant cancer therapy such as surgery, chemotherapy, radiotherapy, thermotherapy, and laser therapy, and may provide a beneficial effect, e.g. reducing tumor size, slowing rate of tumor growth, inhibiting metastasis, or otherwise improving overall clinical condition, without necessarily eradicating the cancer.

The composition can also be administered in therapeutically effective amounts as a portion of an anti-cancer cocktail. An anti-cancer cocktail is a mixture of the polypeptide or modulator of the invention with one or more anti-cancer drugs in addition to a pharmaceutically acceptable carrier for delivery. The use of anti-cancer cocktails as a cancer treatment is routine.

Anti-cancer drugs that are well known in the art and can be used as a treatment in combination with the polypeptide or modulator of the invention include: Actinomycin D, Aminoglutethimide, Asparaginase, Bleomycin, Busulfan, Carboplatin, Carmustine, Chlorambucil, Cisplatin (cis-DDP), Cyclophosphamide, Cytarabine HCl (Cytosine arabinoside), Dacarbazine, Dactinomycin, Daunorubicin HCl, Doxorubicin HCl, Estramustine phosphate sodium, Etoposide (V16-213), Floxuridine, 5-Fluorouracil (5-Fu), Flutamide, Hydroxyurea (hydroxycarbamide), Ifosfamide, Interferon Alpha-2a, Interferon Alpha-2b, Leuprolide acetate (LHRH-releasing factor analog), Lomustine, Mechlorethamine HCl (nitrogen'mustard), Melphalan, Mercaptopurine, Mesna, Methotrexate (MTX), Mitomycin, Mitoxantrone HCl, Octreotide, Plicamycin, Procarbazine HCl, Streptozocin, Tamoxifen citrate, Thioguanine, Thiotepa, Vinblastine sulfate, Vincristine sulfate, Amsacrine, Azacitidine, Hexamethylmelamine, Interleukin-2, Mitoguazone, Pentostatin, Semustine, Teniposide, and Vindesine sulfate.

In addition, therapeutic compositions of the invention may be used for prophylactic treatment of cancer. There are hereditary conditions and/or environmental situations (e.g. exposure to carcinogens) known in the art that predispose an individual to developing cancers. Under these circumstances, it may be beneficial to treat these individuals with therapeutically effective doses of the polypeptide of the invention to reduce the risk of developing cancers.

In vitro models can be used to determine the effective doses of the polypeptide of the invention as a potential cancer treatment. These *in vitro* models include proliferation assays of cultured tumor cells, growth of cultured tumor cells in soft agar (see Freshney, (1987) Culture of Animal Cells: A Manual of Basic Technique, Wily-Liss, New York, NY Ch 18 and Ch 21), tumor systems in nude mice as described in Giovanella et al., J. Natl. Can. Inst., 52: 921-30 (1974), mobility and invasive potential of tumor cells in Boyden Chamber assays as described in Pilkington et al., Anticancer Res., 17: 4107-9 (1997), and angiogenesis assays such as induction of vascularization of the chick chorioallantoic membrane or induction of vascular endothelial cell migration as described in Ribatta et al., Intl. J. Dev. Biol., 40: 1189-97 (1999) and Li et al., Clin. Exp. Metastasis, 17:423-9 (1999), respectively. Suitable tumor cells lines are available, *e.g.* from American Type Tissue Culture Collection catalogs.

4.10.12 RECEPTOR/LIGAND ACTIVITY

10

15

20

25

30

35

A polypeptide of the present invention may also demonstrate activity as receptor, receptor ligand or inhibitor or agonist of receptor/ligand interactions. A polynucleotide of the invention can encode a polypeptide exhibiting such characteristics. Examples of such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and their ligands, receptor phosphatases and their ligands, receptors involved in cell-cell interactions

and their ligands (including without limitation, cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen recognition and development of cellular and humoral immune responses. Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. A protein of the present invention (including, without limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions.

The activity of a polypeptide of the invention may, among other means, be measured by the following methods:

Suitable assays for receptor-ligand activity include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley- Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions 7.28.1-7.28.22), Takai et al., Proc. Natl. Acad. Sci. USA 84:6864-6868, 1987; Bierer et al., J. Exp. Med. 168:1145-1156, 1988; Rosenstein et al., J. Exp. Med. 169:149-160 1989; Stoltenborg et al., J. Immunol. Methods 175:59-68, 1994; Stitt et al., Cell 80:661-670, 1995.

By way of example, the polypeptides of the invention may be used as a receptor for a ligand(s) thereby transmitting the biological activity of that ligand(s). Ligands may be identified through binding assays, affinity chromatography, dihybrid screening assays, BIAcore assays, gel overlay assays, or other methods known in the art.

Studies characterizing drugs or proteins as agonist or antagonist or partial agonists or a partial antagonist require the use of other proteins as competing ligands. The polypeptides of the present invention or ligand(s) thereof may be labeled by being coupled to radioisotopes, colorimetric molecules or a toxin molecules by conventional methods. ("Guide to Protein Purification" Murray P. Deutscher (ed) Methods in Enzymology Vol. 182 (1990) Academic Press, Inc. San Diego). Examples of radioisotopes include, but are not limited to, tritium and carbon-14. Examples of colorimetric molecules include, but are not limited to, fluorescent molecules such as fluorescamine, or rhodamine or other colorimetric molecules. Examples of toxins include, but are not limited, to ricin.

4.10.13 DRUG SCREENING

5

10

15

20

25

30

35

This invention is particularly useful for screening chemical compounds by using the novel polypeptides or binding fragments thereof in any of a variety of drug screening techniques. The polypeptides or fragments employed in such a test may either be free in solution, affixed to a solid support, borne on a cell surface or located intracellularly. One method of drug screening

utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant nucleic acids expressing the polypeptide or a fragment thereof. Drugs are screened against such transformed cells in competitive binding assays. Such cells, either in viable or fixed form, can be used for standard binding assays. One may measure, for example, the formation of complexes between polypeptides of the invention or fragments and the agent being tested or examine the diminution in complex formation between the novel polypeptides and an appropriate cell line, which are well known in the art.

5

10

15

20

25

30

35

Sources for test compounds that may be screened for ability to bind to or modulate (i.e., increase or decrease) the activity of polypeptides of the invention include (1) inorganic and organic chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of either random or mimetic peptides, oligonucleotides or organic molecules.

Chemical libraries may be readily synthesized or purchased from a number of commercial sources, and may include structural analogs of known compounds or compounds that are identified as "hits" or "leads" via natural product screening.

The sources of natural product libraries are microorganisms (including bacteria and fungi), animals, plants or other vegetation, or marine organisms, and libraries of mixtures for screening may be created by: (1) fermentation and extraction of broths from soil, plant or marine microorganisms or (2) extraction of the organisms themselves. Natural product libraries include polyketides, non-ribosomal peptides, and (non-naturally occurring) variants thereof. For a review, see *Science* 282:63-68 (1998).

Combinatorial libraries are composed of large numbers of peptides, oligonucleotides or organic compounds and can be readily prepared by traditional automated synthesis methods, PCR, cloning or proprietary synthetic methods. Of particular interest are peptide and oligonucleotide combinatorial libraries. Still other libraries of interest include peptide, protein, peptidomimetic, multiparallel synthetic collection, recombinatorial, and polypeptide libraries. For a review of combinatorial chemistry and libraries created therefrom, see Myers, Curr. Opin. Biotechnol. 8:701-707 (1997). For reviews and examples of peptidomimetic libraries, see Al-Obeidi et al., Mol. Biotechnol, 9(3):205-23 (1998); Hruby et al., Curr Opin Chem Biol, 1(1):114-19 (1997); Dorner et al., Bioorg Med Chem, 4(5):709-15 (1996) (alkylated dipeptides).

Identification of modulators through use of the various libraries described herein permits modification of the candidate "hit" (or "lead") to optimize the capacity of the "hit" to bind a polypeptide of the invention. The molecules identified in the binding assay are then tested for antagonist or agonist activity in *in vivo* tissue culture or animal models that are well known in the art. In brief, the molecules are titrated into a plurality of cell cultures or animals and then tested for either cell/animal death or prolonged survival of the animal/cells.

The binding molecules thus identified may be complexed with toxins, e.g., ricin or cholera, or with other compounds that are toxic to cells such as radioisotopes. The toxin-binding molecule complex is then targeted to a tumor or other cell by the specificity of the binding molecule for a polypeptide of the invention. Alternatively, the binding molecules may be complexed with imaging agents for targeting and imaging purposes.

4.10.14 ASSAY FOR RECEPTOR ACTIVITY

The invention also provides methods to detect specific binding of a polypeptide e.g. a ligand or a receptor. The art provides numerous assays particularly useful for identifying previously unknown binding partners for receptor polypeptides of the invention. For example, expression cloning using mammalian or bacterial cells, or dihybrid screening assays can be used to identify polynucleotides encoding binding partners. As another example, affinity chromatography with the appropriate immobilized polypeptide of the invention can be used to isolate polypeptides that recognize and bind polypeptides of the invention. There are a number of different libraries used for the identification of compounds, and in particular small molecules, that modulate (i.e., increase or decrease) biological activity of a polypeptide of the invention. Ligands for receptor polypeptides of the invention can also be identified by adding exogenous ligands, or cocktails of ligands to two cells populations that are genetically identical except for the expression of the receptor of the invention: one cell population expresses the receptor of the invention whereas the other does not. The response of the two cell populations to the addition of ligands(s) are then compared. Alternatively, an expression library can be co-expressed with the polypeptide of the invention in cells and assayed for an autocrine response to identify potential ligand(s). As still another example, BIAcore assays, gel overlay assays, or other methods known in the art can be used to identify binding partner polypeptides, including, (1) organic and inorganic chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of random peptides, oligonucleotides or organic molecules.

The role of downstream intracellular signaling molecules in the signaling cascade of the polypeptide of the invention can be determined. For example, a chimeric protein in which the cytoplasmic domain of the polypeptide of the invention is fused to the extracellular portion of a protein, whose ligand has been identified, is produced in a host cell. The cell is then incubated with the ligand specific for the extracellular portion of the chimeric protein, thereby activating the chimeric receptor. Known downstream proteins involved in intracellular signaling can then be assayed for expected modifications *i.e.* phosphorylation. Other methods known to those in the art can also be used to identify signaling molecules involved in receptor activity.

5

10

15

20

25

30

4.10.15 ANTI-INFLAMMATORY ACTIVITY

Compositions of the present invention may also exhibit anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production of other factors which more directly inhibit or promote an inflammatory response. Compositions with such activities can be used to treat inflammatory conditions including chronic or acute conditions), including without limitation intimation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from over production of cytokines such as TNF or IL-1. Compositions of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material. Compositions of this invention may be utilized to prevent or treat conditions such as, but not limited to, sepsis, acute pancreatitis, endotoxin shock, cytokine induced shock, rheumatoid arthritis, chronic inflammatory arthritis, pancreatic cell damage from diabetes mellitus type 1. graft versus host disease, inflammatory bowel disease, inflamation associated with pulmonary disease, other autoimmune disease or inflammatory disease, an antiproliferative agent such as for acute or chronic mylegenous leukemia or in the prevention of premature labor secondary to intrauterine infections.

4.10.16 LEUKEMIAS

5

10

15

20

35

Leukemias and related disorders may be treated or prevented by administration of a

therapeutic that promotes or inhibits function of the polynucleotides and/or polypeptides of the invention. Such leukemias and related disorders include but are not limited to acute leukemia, acute lymphocytic leukemia, acute myelocytic leukemia, myeloblastic, promyelocytic, myelomonocytic, monocytic, erythroleukemia, chronic leukemia, chronic myelocytic (granulocytic) leukemia and chronic lymphocytic leukemia (for a review of such disorders, see

Fishman et al., 1985, Medicine, 2d Ed., J.B. Lippincott Co., Philadelphia).

4.10.17 NERVOUS SYSTEM DISORDERS

Nervous system disorders, involving cell types which can be tested for efficacy of intervention with compounds that modulate the activity of the polynucleotides and/or polypeptides of the invention, and which can be treated upon thus observing an indication of

therapeutic utility, include but are not limited to nervous system injuries, and diseases or disorders which result in either a disconnection of axons, a diminution or degeneration of neurons, or demyelination. Nervous system lesions which may be treated in a patient (including human and non-human mammalian patients) according to the invention include but are not limited to the following lesions of either the central (including spinal cord, brain) or peripheral nervous systems:

5

10

15

20

25

30

- (i) traumatic lesions, including lesions caused by physical injury or associated with surgery, for example, lesions which sever a portion of the nervous system, or compression injuries;
- (ii) ischemic lesions, in which a lack of oxygen in a portion of the nervous system results in neuronal injury or death, including cerebral infarction or ischemia, or spinal cord infarction or ischemia;
- (iii) infectious lesions, in which a portion of the nervous system is destroyed or injured as a result of infection, for example, by an abscess or associated with infection by human immunodeficiency virus, herpes zoster, or herpes simplex virus or with Lyme disease, tuberculosis, syphilis;
- (iv) degenerative lesions, in which a portion of the nervous system is destroyed or injured as a result of a degenerative process including but not limited to degeneration associated with Parkinson's disease, Alzheimer's disease, Huntington's chorea, or amyotrophic lateral sclerosis;
- (v) lesions associated with nutritional diseases or disorders, in which a portion of the nervous system is destroyed or injured by a nutritional disorder or disorder of metabolism including but not limited to, vitamin B12 deficiency, folic acid deficiency, Wernicke disease, tobacco-alcohol amblyopia, Marchiafava-Bignami disease (primary degeneration of the corpus callosum), and alcoholic cerebellar degeneration;
- (vi) neurological lesions associated with systemic diseases including but not limited to diabetes (diabetic neuropathy, Bell's palsy), systemic lupus erythematosus, carcinoma, or sarcoidosis;
- (vii) lesions caused by toxic substances including alcohol, lead, or particular neurotoxins; and
- (viii) demyelinated lesions in which a portion of the nervous system is destroyed or injured by a demyelinating disease including but not limited to multiple sclerosis, human immunodeficiency virus-associated myelopathy, transverse myelopathy or various etiologies, progressive multifocal leukoencephalopathy, and central pontine myelinolysis.

Therapeutics which are useful according to the invention for treatment of a nervous system disorder may be selected by testing for biological activity in promoting the survival or differentiation of neurons. For example, and not by way of limitation, therapeutics which elicit any of the following effects may be useful according to the invention:

(i) increased survival time of neurons in culture;

5

10

15

20

25

- (ii) increased sprouting of neurons in culture or in vivo;
- (iii) increased production of a neuron-associated molecule in culture or *in vivo*, *e.g.*, choline acetyltransferase or acetylcholinesterase with respect to motor neurons; or
 - (iv) decreased symptoms of neuron dysfunction in vivo.

Such effects may be measured by any method known in the art. In preferred, non-limiting embodiments, increased survival of neurons may be measured by the method set forth in Arakawa et al. (1990, J. Neurosci. 10:3507-3515); increased sprouting of neurons may be detected by methods set forth in Pestronk et al. (1980, Exp. Neurol. 70:65-82) or Brown et al. (1981, Ann. Rev. Neurosci. 4:17-42); increased production of neuron-associated molecules may be measured by bioassay, enzymatic assay, antibody binding, Northern blot assay, etc., depending on the molecule to be measured; and motor neuron dysfunction may be measured by assessing the physical manifestation of motor neuron disorder, e.g., weakness, motor neuron conduction velocity, or functional disability.

In specific embodiments, motor neuron disorders that may be treated according to the invention include but are not limited to disorders such as infarction, infection, exposure to toxin, trauma, surgical damage, degenerative disease or malignancy that may affect motor neurons as well as other components of the nervous system, as well as disorders that selectively affect neurons such as amyotrophic lateral sclerosis, and including but not limited to progressive spinal muscular atrophy, progressive bulbar palsy, primary lateral sclerosis, infantile and juvenile muscular atrophy, progressive bulbar paralysis of childhood (Fazio-Londe syndrome), poliomyelitis and the post polio syndrome, and Hereditary Motorsensory Neuropathy (Charcot-Marie-Tooth Disease).

4.10.18 OTHER ACTIVITIES

A polypeptide of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape

(such as, for example, breast augmentation or diminution, change in bone form or shape);

effecting biorhythms or circadian cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, co-factors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

15 4.10.19 IDENTIFICATION OF POLYMORPHISMS

5

10

20

25

30

35

The demonstration of polymorphisms makes possible the identification of such polymorphisms in human subjects and the pharmacogenetic use of this information for diagnosis and treatment. Such polymorphisms may be associated with, e.g., differential predisposition or susceptibility to various disease states (such as disorders involving inflammation or immune response) or a differential response to drug administration, and this genetic information can be used to tailor preventive or therapeutic treatment appropriately. For example, the existence of a polymorphism associated with a predisposition to inflammation or autoimmune disease makes possible the diagnosis of this condition in humans by identifying the presence of the polymorphism.

Polymorphisms can be identified in a variety of ways known in the art which all generally involve obtaining a sample from a patient, analyzing DNA from the sample, optionally involving isolation or amplification of the DNA, and identifying the presence of the polymorphism in the DNA. For example, PCR may be used to amplify an appropriate fragment of genomic DNA which may then be sequenced. Alternatively, the DNA may be subjected to allele-specific oligonucleotide hybridization (in which appropriate oligonucleotides are hybridized to the DNA under conditions permitting detection of a single base mismatch) or to a single nucleotide extension assay (in which an oligonucleotide that hybridizes immediately adjacent to the position of the polymorphism is extended with one or more labeled nucleotides). In addition, traditional restriction fragment length polymorphism analysis (using restriction enzymes that provide differential digestion of the genomic DNA depending on the presence or

absence of the polymorphism) may be performed. Arrays with nucleotide sequences of the present invention can be used to detect polymorphisms. The array can comprise modified nucleotide sequences of the present invention in order to detect the nucleotide sequences of the present invention. In the alternative, any one of the nucleotide sequences of the present invention can be placed on the array to detect changes from those sequences.

Alternatively a polymorphism resulting in a change in the amino acid sequence could also be detected by detecting a corresponding change in amino acid sequence of the protein, e.g., by an antibody specific to the variant sequence.

4.10.20 ARTHRITIS AND INFLAMMATION

5

10

15

20

25

30

The immunosuppressive effects of the compositions of the invention against rheumatoid arthritis is determined in an experimental animal model system. The experimental model system is adjuvant induced arthritis in rats, and the protocol is described by J. Holoshitz, et at., 1983, Science, 219:56, or by B. Waksman et al., 1963, Int. Arch. Allergy Appl. Immunol., 23:129. Induction of the disease can be caused by a single injection, generally intradermally, of a suspension of killed Mycobacterium tuberculosis in complete Freund's adjuvant (CFA). The route of injection can vary, but rats may be injected at the base of the tail with an adjuvant mixture. The polypeptide is administered in phosphate buffered solution (PBS) at a dose of about 1-5 mg/kg. The control consists of administering PBS only.

The procedure for testing the effects of the test compound would consist of intradermally injecting killed Mycobacterium tuberculosis in CFA followed by immediately administering the test compound and subsequent treatment every other day until day 24. At 14, 15, 18, 20, 22, and 24 days after injection of Mycobacterium CFA, an overall arthritis score may be obtained as described by J. Holoskitz above. An analysis of the data would reveal that the test compound would have a dramatic affect on the swelling of the joints as measured by a decrease of the arthritis score.

4.11 THERAPEUTIC METHODS

The compositions (including polypeptide fragments, analogs, variants and antibodies or other binding partners or modulators including antisense polynucleotides) of the invention have numerous applications in a variety of therapeutic methods. Examples of therapeutic applications include, but are not limited to, those exemplified herein.

4.11.1 EXAMPLE

One embodiment of the invention is the administration of an effective amount of the polypeptides or other composition of the invention to individuals affected by a disease or disorder that can be modulated by regulating the peptides of the invention. While the mode of administration is not particularly important, parenteral administration is preferred. An exemplary mode of administration is to deliver an intravenous bolus. The dosage of the polypeptides or other composition of the invention will normally be determined by the prescribing physician. It is to be expected that the dosage will vary according to the age, weight, condition and response of the individual patient. Typically, the amount of polypeptide administered per dose will be in the range of about 0.01µg/kg to 100 mg/kg of body weight, with the preferred dose being about 0.1µg/kg to 10 mg/kg of patient body weight. For parenteral administration, polypeptides of the invention will be formulated in an injectable form combined with a pharmaceutically acceptable parenteral vehicle. Such vehicles are well known in the art and examples include water, saline, Ringer's solution, dextrose solution, and solutions consisting of small amounts of the human serum albumin. The vehicle may contain minor amounts of additives that maintain the isotonicity and stability of the polypeptide or other active ingredient. The preparation of such solutions is within the skill of the art.

4.12 PHARMACEUTICAL FORMULATIONS AND ROUTES OF ADMINISTRATION

5

10

15

20

25

30

35

A protein or other composition of the present invention (from whatever source derived, including without limitation from recombinant and non-recombinant sources and including antibodies and other binding partners of the polypeptides of the invention) may be administered to a patient in need, by itself, or in pharmaceutical compositions where it is mixed with suitable carriers or excipient(s) at doses to treat or ameliorate a variety of disorders. Such a composition may optionally contain (in addition to protein or other active ingredient and a carrier) diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials well known in the art. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredient(s). The characteristics of the carrier will depend on the route of administration. The pharmaceutical composition of the invention may also contain cytokines, lymphokines, or other hematopoietic factors such as M-CSF, GM-CSF, TNF, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IFN, TNF0, TNF1, TNF2, G-CSF, Meg-CSF, thrombopoietin, stem cell factor, and erythropoietin. In further compositions, proteins of the invention may be combined with other agents beneficial to the treatment of the disease or disorder in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet-derived growth

factor (PDGF), transforming growth factors (TGF- α and TGF- β), insulin-like growth factor (IGF), as well as cytokines described herein.

The pharmaceutical composition may further contain other agents which either enhance the activity of the protein or other active ingredient or complement its activity or use in treatment. Such additional factors and/or agents may be included in the pharmaceutical composition to produce a synergistic effect with protein or other active ingredient of the invention, or to minimize side effects. Conversely, protein or other active ingredient of the present invention may be included in formulations of the particular clotting factor, cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent to minimize side effects of the clotting factor, cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent (such as IL-1Ra, IL-1 Hy1, IL-1 Hy2, anti-TNF, corticosteroids, immunosuppressive agents). A protein of the present invention may be active in multimers (e.g., heterodimers or homodimers) or complexes with itself or other proteins. As a result, pharmaceutical compositions of the invention may comprise a protein of the invention in such multimeric or complexed form.

10

15

30

35

As an alternative to being included in a pharmaceutical composition of the invention including a first protein, a second protein or a therapeutic agent may be concurrently administered with the first protein (e.g., at the same time, or at differing times provided that therapeutic concentrations of the combination of agents is achieved at the treatment site). 20 Techniques for formulation and administration of the compounds of the instant application may be found in "Remington's Pharmaceutical Sciences," Mack Publishing Co., Easton, PA, latest edition. A therapeutically effective dose further refers to that amount of the compound sufficient to result in amelioration of symptoms, e.g., treatment, healing, prevention or amelioration of the relevant medical condition, or an increase in rate of treatment, healing, prevention or 25 amelioration of such conditions. When applied to an individual active ingredient, administered alone, a therapeutically effective dose refers to that ingredient alone. When applied to a combination, a therapeutically effective dose refers to combined amounts of the active ingredients that result in the therapeutic effect, whether administered in combination, serially or simultaneously.

In practicing the method of treatment or use of the present invention, a therapeutically effective amount of protein or other active ingredient of the present invention is administered to a mammal having a condition to be treated. Protein or other active ingredient of the present invention may be administered in accordance with the method of the invention either alone or in combination with other therapies such as treatments employing cytokines, lymphokines or other hematopoietic factors. When co- administered with one or more cytokines, lymphokines or other

hematopoietic factors, protein or other active ingredient of the present invention may be administered either simultaneously with the cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors, or sequentially. If administered sequentially, the attending physician will decide on the appropriate sequence of administering protein or other active ingredient of the present invention in combination with cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors.

4.12.1 ROUTES OF ADMINISTRATION

5

10

15

20

25

30

Suitable routes of administration may, for example, include oral, rectal, transmucosal, or intestinal administration; parenteral delivery, including intramuscular, subcutaneous, intramedullary injections, as well as intrathecal, direct intraventricular, intravenous, intraperitoneal, intranasal, or intraocular injections. Administration of protein or other active ingredient of the present invention used in the pharmaceutical composition or to practice the method of the present invention can be carried out in a variety of conventional ways, such as oral ingestion, inhalation, topical application or cutaneous, subcutaneous, intraperitoneal, parenteral or intravenous injection. Intravenous administration to the patient is preferred.

Alternately, one may administer the compound in a local rather than systemic manner, for example, via injection of the compound directly into a arthritic joints or in fibrotic tissue, often in a depot or sustained release formulation. In order to prevent the scarring process frequently occurring as complication of glaucoma surgery, the compounds may be administered topically, for example, as eye drops. Furthermore, one may administer the drug in a targeted drug delivery system, for example, in a liposome coated with a specific antibody, targeting, for example, arthritic or fibrotic tissue. The liposomes will be targeted to and taken up selectively by the afflicted tissue.

The polypeptides of the invention are administered by any route that delivers an effective dosage to the desired site of action. The determination of a suitable route of administration and an effective dosage for a particular indication is within the level of skill in the art. Preferably for wound treatment, one administers the therapeutic compound directly to the site. Suitable dosage ranges for the polypeptides of the invention can be extrapolated from these dosages or from similar studies in appropriate animal models. Dosages can then be adjusted as necessary by the clinician to provide maximal therapeutic benefit.

4.12.2 COMPOSITIONS/FORMULATIONS

Pharmaceutical compositions for use in accordance with the present invention thus may be formulated in a conventional manner using one or more physiologically acceptable carriers

5

10

15

20

25

30

35

comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. These pharmaceutical compositions may be manufactured in a manner that is itself known, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping or lyophilizing processes. Proper formulation is dependent upon the route of administration chosen. When a therapeutically effective amount of protein or other active ingredient of the present invention is administered orally, protein or other active ingredient of the present invention will be in the form of a tablet, capsule, powder, solution or elixir. When administered in tablet form, the pharmaceutical composition of the invention may additionally contain a solid carrier such as a gelatin or an adjuvant. The tablet, capsule, and powder contain from about 5 to 95% protein or other active ingredient of the present invention, and preferably from about 25 to 90% protein or other active ingredient of the present invention. When administered in liquid form, a liquid carrier such as water, petroleum, oils of animal or plant origin such as peanut oil, mineral oil, soybean oil, or sesame oil, or synthetic oils may be added. The liquid form of the pharmaceutical composition may further contain physiological saline solution, dextrose or other saccharide solution, or glycols such as ethylene glycol, propylene glycol or polyethylene glycol. When administered in liquid form, the pharmaceutical composition contains from about 0.5 to 90% by weight of protein or other active ingredient of the present invention, and preferably from about 1 to 50% protein or other active ingredient of the present invention.

When a therapeutically effective amount of protein or other active ingredient of the present invention is administered by intravenous, cutaneous or subcutaneous injection, protein or other active ingredient of the present invention will be in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such parenterally acceptable protein or other active ingredient solutions, having due regard to pH, isotonicity, stability, and the like, is within the skill in the art. A preferred pharmaceutical composition for intravenous, cutaneous, or subcutaneous injection should contain, in addition to protein or other active ingredient of the present invention, an isotonic vehicle such as Sodium Chloride Injection, Ringer's Injection, Dextrose Injection, Dextrose and Sodium Chloride Injection, Lactated Ringer's Injection, or other vehicle as known in the art. The pharmaceutical composition of the present invention may also contain stabilizers, preservatives, buffers, antioxidants, or other additives known to those of skill in the art. For injection, the agents of the invention may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks's solution, Ringer's solution, or physiological saline buffer. For transmucosal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

5

10

15

20

25

30

35

For oral administration, the compounds can be formulated readily by combining the active compounds with pharmaceutically acceptable carriers well known in the art. Such carriers enable the compounds of the invention to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions and the like, for oral ingestion by a patient to be treated. Pharmaceutical preparations for oral use can be obtained from a solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are, in particular, fillers such as sugars, including lactose, sucrose, mannitol, or sorbitol; cellulose preparations such as, for example, maize starch, wheat starch, rice starch, potato starch, gelatin, gum tragacanth, methyl cellulose, hydroxypropylmethyl-cellulose, sodium carboxymethylcellulose, and/or polyvinylpyrrolidone (PVP). If desired, disintegrating agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, or alginic acid or a salt thereof such as sodium alginate. Dragee cores are provided with suitable coatings. For this purpose, concentrated sugar solutions may be used, which may optionally contain gum arabic, talc, polyvinyl pyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for identification or to characterize different combinations of active compound doses.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a plasticizer, such as glycerol or sorbitol. The push-fit capsules can contain the active ingredients in admixture with filler such as lactose, binders such as starches, and/or lubricants such as talc or magnesium stearate and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid paraffin, or liquid polyethylene glycols. In addition, stabilizers may be added. All formulations for oral administration should be in dosages suitable for such administration. For buccal administration, the compositions may take the form of tablets or lozenges formulated in conventional manner.

For administration by inhalation, the compounds for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebuliser, with the use of a suitable propellant, e.g., dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of, e.g., gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch. The compounds may be formulated for parenteral

administration by injection, e.g., by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, e.g., in ampules or in multi-dose containers, with an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents.

Pharmaceutical formulations for parenteral administration include aqueous solutions of the active compounds in water-soluble form. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions. Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, e.g., sterile pyrogen-free water, before use.

The compounds may also be formulated in rectal compositions such as suppositories or retention enemas, e.g., containing conventional suppository bases such as cocoa butter or other glycerides. In addition to the formulations described previously, the compounds may also be formulated as a depot preparation. Such long acting formulations may be administered by implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for example, the compounds may be formulated with suitable polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt.

A pharmaceutical carrier for the hydrophobic compounds of the invention is a co-solvent system comprising benzyl alcohol, a nonpolar surfactant, a water-miscible organic polymer, and an aqueous phase. The co-solvent system may be the VPD co-solvent system. VPD is a solution of 3% w/v benzyl alcohol, 8% w/v of the nonpolar surfactant polysorbate 80, and 65% w/v polyethylene glycol 300, made up to volume in absolute ethanol. The VPD co-solvent system (VPD:5W) consists of VPD diluted 1:1 with a 5% dextrose in water solution. This co-solvent system dissolves hydrophobic compounds well, and itself produces low toxicity upon systemic administration. Naturally, the proportions of a co-solvent system may be varied considerably without destroying its solubility and toxicity characteristics. Furthermore, the identity of the co-solvent components may be varied: for example, other low-toxicity nonpolar surfactants may be used instead of polysorbate 80; the fraction size of polyethylene glycol may be varied; other biocompatible polymers may replace polyethylene glycol, e.g. polyvinyl pyrrolidone; and other

sugars or polysaccharides may substitute for dextrose. Alternatively, other delivery systems for hydrophobic pharmaceutical compounds may be employed. Liposomes and emulsions are well known examples of delivery vehicles or carriers for hydrophobic drugs. Certain organic solvents such as dimethylsulfoxide also may be employed, although usually at the cost of greater toxicity. Additionally, the compounds may be delivered using a sustained-release system, such as semipermeable matrices of solid hydrophobic polymers containing the therapeutic agent. Various types of sustained-release materials have been established and are well known by those skilled in the art. Sustained-release capsules may, depending on their chemical nature, release the compounds for a few weeks up to over 100 days. Depending on the chemical nature and the biological stability of the therapeutic reagent, additional strategies for protein or other active ingredient stabilization may be employed.

The pharmaceutical compositions also may comprise suitable solid or gel phase carriers or excipients. Examples of such carriers or excipients include but are not limited to calcium carbonate, calcium phosphate, various sugars, starches, cellulose derivatives, gelatin, and polymers such as polyethylene glycols. Many of the active ingredients of the invention may be provided as salts with pharmaceutically compatible counter ions. Such pharmaceutically acceptable base addition salts are those salts which retain the biological effectiveness and properties of the free acids and which are obtained by reaction with inorganic or organic bases such as sodium hydroxide, magnesium hydroxide, ammonia, trialkylamine, dialkylamine, monoalkylamine, dibasic amino acids, sodium acetate, potassium benzoate, triethanol amine and the like.

The pharmaceutical composition of the invention may be in the form of a complex of the protein(s) or other active ingredient(s) of present invention along with protein or peptide antigens. The protein and/or peptide antigen will deliver a stimulatory signal to both B and T lymphocytes. B lymphocytes will respond to antigen through their surface immunoglobulin receptor. T lymphocytes will respond to antigen through the T cell receptor (TCR) following presentation of the antigen by MHC proteins. MHC and structurally related proteins including those encoded by class I and class II MHC genes on host cells will serve to present the peptide antigen(s) to T lymphocytes. The antigen components could also be supplied as purified MHC-peptide complexes alone or with co-stimulatory molecules that can directly signal T cells. Alternatively antibodies able to bind surface immunoglobulin and other molecules on B cells as well as antibodies able to bind the TCR and other molecules on T cells can be combined with the pharmaceutical composition of the invention.

The pharmaceutical composition of the invention may be in the form of a liposome in which protein of the present invention is combined, in addition to other pharmaceutically

acceptable carriers, with amphipathic agents such as lipids which exist in aggregated form as micelles, insoluble monolayers, liquid crystals, or lamellar layers in aqueous solution. Suitable lipids for liposomal formulation include, without limitation, monoglycerides, diglycerides, sulfatides, lysolecithins, phospholipids, saponin, bile acids, and the like. Preparation of such liposomal formulations is within the level of skill in the art, as disclosed, for example, in U.S. Patent Nos. 4,235,871; 4,501,728; 4,837,028; and 4,737,323, all of which are incorporated herein by reference.

5

10

15

20

25

30

The amount of protein or other active ingredient of the present invention in the pharmaceutical composition of the present invention will depend upon the nature and severity of the condition being treated, and on the nature of prior treatments which the patient has undergone. Ultimately, the attending physician will decide the amount of protein or other active ingredient of the present invention with which to treat each individual patient. Initially, the attending physician will administer low doses of protein or other active ingredient of the present invention and observe the patient's response. Larger doses of protein or other active ingredient of the present invention may be administered until the optimal therapeutic effect is obtained for the patient, and at that point the dosage is not increased further. It is contemplated that the various pharmaceutical compositions used to practice the method of the present invention should contain about 0.01 µg to about 100 mg (preferably about 0.1 µg to about 10 mg, more preferably about 0.1 µg to about 1 mg) of protein or other active ingredient of the present invention per kg body weight. For compositions of the present invention which are useful for bone, cartilage, tendon or ligament regeneration, the therapeutic method-includes administering the composition topically, systematically, or locally as an implant or device. When administered, the therapeutic composition for use in this invention is, of course, in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably be encapsulated or injected in a viscous form for delivery to the site of bone, cartilage or tissue damage. Topical administration may be suitable for wound healing and tissue repair. Therapeutically useful agents other than a protein or other active ingredient of the invention which may also optionally be included in the composition as described above, may alternatively or additionally, be administered simultaneously or sequentially with the composition in the methods of the invention. Preferably for bone and/or cartilage formation, the composition would include a matrix capable of delivering the protein-containing or other active ingredient-containing composition to the site of bone and/or cartilage damage, providing a structure for the developing bone and cartilage and optimally capable of being resorbed into the body. Such matrices may be formed of materials presently in use for other implanted medical applications.

The choice of matrix material is based on biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. The particular application of the compositions will define the appropriate formulation. Potential matrices for the compositions may be biodegradable and chemically defined calcium sulfate, tricalcium phosphate, hydroxyapatite, polylactic acid, polyglycolic acid and polyanhydrides. Other potential materials are biodegradable and biologically well-defined, such as bone or dermal collagen. Further matrices are comprised of pure proteins or extracellular matrix components. Other potential matrices are nonbiodegradable and chemically defined, such as sintered hydroxyapatite, bioglass, aluminates, or other ceramics. Matrices may be comprised of combinations of any of the above mentioned types of material, such as polylactic acid and hydroxyapatite or collagen and tricalcium phosphate. The bioceramics may be altered in composition, such as in calcium-aluminate-phosphate and processing to alter pore size, particle size, particle shape, and biodegradability. Presently preferred is a 50:50 (mole weight) copolymer of lactic acid and glycolic acid in the form of porous particles having diameters ranging from 150 to 800 microns. In some applications, it will be useful to utilize a sequestering agent, such as carboxymethyl cellulose or autologous blood clot, to prevent the protein compositions from disassociating from the matrix.

5

10

15

A preferred family of sequestering agents is cellulosic materials such as alkylcelluloses (including hydroxyalkylcelluloses), including methylcellulose, ethylcellulose, 20 hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropyl-methylcellulose, and carboxymethylcellulose, the most preferred being cationic salts of carboxymethylcellulose (CMC). Other preferred sequestering agents include hyaluronic acid, sodium alginate, poly(ethylene glycol), polyoxyethylene oxide, carboxyvinyl polymer and poly(vinyl alcohol). The amount of sequestering agent useful herein is 0.5-20 wt %, preferably 1-10 wt % based on 25 total formulation weight, which represents the amount necessary to prevent desorption of the protein from the polymer matrix and to provide appropriate handling of the composition, yet not so much that the progenitor cells are prevented from infiltrating the matrix, thereby providing the protein the opportunity to assist the osteogenic activity of the progenitor cells. In further compositions, proteins or other active ingredients of the invention may be combined with other 30 agents beneficial to the treatment of the bone and/or cartilage defect, wound, or tissue in question. These agents include various growth factors such as epidermal growth factor (EGF). platelet derived growth factor (PDGF), transforming growth factors (TGF-α and TGF-β), and insulin-like growth factor (IGF).

The therapeutic compositions are also presently valuable for veterinary applications.

Particularly domestic animals and thoroughbred horses, in addition to humans, are desired

patients for such treatment with proteins or other active ingredients of the present invention. The dosage regimen of a protein-containing pharmaceutical composition to be used in tissue regeneration will be determined by the attending physician considering various factors which modify the action of the proteins, e.g., amount of tissue weight desired to be formed, the site of damage, the condition of the damaged tissue, the size of a wound, type of damaged tissue (e.g., bone), the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and with inclusion of other proteins in the pharmaceutical composition. For example, the addition of other known growth factors, such as IGF I (insulin like growth factor I), to the final composition, may also effect the dosage. Progress can be monitored by periodic assessment of tissue/bone growth and/or repair, for example, X-rays, histomorphometric determinations and tetracycline labeling.

Polynucleotides of the present invention can also be used for gene therapy. Such polynucleotides can be introduced either in vivo or ex vivo into cells for expression in a mammalian subject. Polynucleotides of the invention may also be administered by other known methods for introduction of nucleic acid into a cell or organism (including, without limitation, in the form of viral vectors or naked DNA). Cells may also be cultured ex vivo in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced in vivo for therapeutic purposes.

20

25

30

15

5

10

4.12.3 EFFECTIVE DOSAGE

Pharmaceutical compositions suitable for use in the present invention include compositions wherein the active ingredients are contained in an effective amount to achieve its intended purpose. More specifically, a therapeutically effective amount means an amount effective to prevent development of or to alleviate the existing symptoms of the subject being treated. Determination of the effective amount is well within the capability of those skilled in the art, especially in light of the detailed disclosure provided herein. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from appropriate in vitro assays. For example, a dose can be formulated in animal models to achieve a circulating concentration range that can be used to more accurately determine useful doses in humans. For example, a dose can be formulated in animal models to achieve a circulating concentration range that includes the IC₅₀ as determined in cell culture (*i.e.*, the concentration of the test compound which achieves a half-maximal inhibition of the protein's biological activity). Such information can be used to more accurately determine useful doses in humans.

A therapeutically effective dose refers to that amount of the compound that results in amelioration of symptoms or a prolongation of survival in a patient. Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., for determining the LD₅₀ (the dose lethal to 50% of the population) and the ED₅₀ (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio between LD₅₀ and ED₅₀. Compounds which exhibit high therapeutic indices are preferred. The data obtained from these cell culture assays and animal studies can be used in formulating a range of dosage for use in human. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED₅₀ with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. The exact formulation, route of administration and dosage can be chosen by the individual physician in view of the patient's condition. See, e.g., Fingl et al., 1975, in "The Pharmacological Basis of Therapeutics", Ch. 1 p.1. Dosage amount and interval may be adjusted individually to provide plasma levels of the active moiety which are sufficient to maintain the desired effects, or minimal effective concentration (MEC). The MEC will vary for each compound but can be estimated from in vitro data. Dosages necessary to achieve the MEC will depend on individual characteristics and route of administration. However, HPLC assays or bioassays can be used to determine plasma concentrations.

Dosage intervals can also be determined using MEC value. Compounds should be administered using a regimen which maintains plasma levels above the MEC for 10-90% of the time, preferably between 30-90% and most preferably between 50-90%. In cases of local administration or selective uptake, the effective local concentration of the drug may not be related to plasma concentration.

An exemplary dosage regimen for polypeptides or other compositions of the invention will be in the range of about 0.01 μ g/kg to 100 mg/kg of body weight daily, with the preferred dose being about 0.1 μ g/kg to 25 mg/kg of patient body weight daily, varying in adults and children. Dosing may be once daily, or equivalent doses may be delivered at longer or shorter intervals.

The amount of composition administered will, of course, be dependent on the subject being treated, on the subject's age and weight, the severity of the affliction, the manner of administration and the judgment of the prescribing physician.

4.12.4 PACKAGING

5

10

15

20

25

30

The compositions may, if desired, be presented in a pack or dispenser device which may contain one or more unit dosage forms containing the active ingredient. The pack may, for example, comprise metal or plastic foil, such as a blister pack. The pack or dispenser device may be accompanied by instructions for administration. Compositions comprising a compound of the invention formulated in a compatible pharmaceutical carrier may also be prepared, placed in an appropriate container, and labeled for treatment of an indicated condition.

4.13 ANTIBODIES

5

10

15

20

25

30

35

Also included in the invention are antibodies to proteins, or fragments of proteins of the invention. The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin (Ig) molecules, *i.e.*, molecules that contain an antigen binding site that specifically binds (immunoreacts with) an antigen. Such antibodies include, but are not limited to, polyclonal, monoclonal, chimeric, single chain, F_{ab} , $F_{ab'}$ and $F_{(ab')2}$ fragments, and an F_{ab} expression library. In general, an antibody molecule obtained from humans relates to any of the classes IgG, IgM, IgA, IgE and IgD, which differ from one another by the nature of the heavy chain present in the molecule. Certain classes have subclasses as well, such as IgG₁, IgG₂, and others. Furthermore, in humans, the light chain may be a kappa chain or a lambda chain. Reference herein to antibodies includes a reference to all such classes, subclasses and types of human antibody species.

An isolated related protein of the invention may be intended to serve as an antigen, or a portion or fragment thereof, and additionally can be used as an immunogen to generate antibodies that immunospecifically bind the antigen, using standard techniques for polyclonal and monoclonal antibody preparation. The full-length protein can be used or, alternatively, the invention provides antigenic peptide fragments of the antigen for use as immunogens. An antigenic peptide fragment comprises at least 6 amino acid residues of the amino acid sequence of the full length protein, (for example the amino acid sequence shown in SEQ ID NO: 1351), and encompasses an epitope thereof such that an antibody raised against the peptide forms a specific immune complex with the full length protein or with any fragment that contains the epitope. Preferably, the antigenic peptide comprises at least 10 amino acid residues, or at least 15 amino acid residues, or at least 20 amino acid residues. Preferred epitopes encompassed by the antigenic peptide are regions of the protein that are located on its surface; commonly these are hydrophilic regions.

In certain embodiments of the invention, at least one epitope encompassed by the antigenic peptide is a region of -related protein that is located on the surface of the protein, e.g., a hydrophilic region. A hydrophobicity analysis of the human related protein sequence will

indicate which regions of a related protein are particularly hydrophilic and, therefore, are likely to encode surface residues useful for targeting antibody production. As a means for targeting antibody production, hydropathy plots showing regions of hydrophilicity and hydrophobicity may be generated by any method well known in the art, including, for example, the Kyte Doolittle or the Hopp Woods methods, either with or without Fourier transformation. See, e.g., Hopp and Woods, 1981, Proc. Nat. Acad. Sci. USA 78: 3824-3828; Kyte and Doolittle 1982, J. Mol. Biol. 157: 105-142, each of which is incorporated herein by reference in its entirety. Antibodies that are specific for one or more domains within an antigenic protein, or derivatives, fragments, analogs or homologs thereof, are also provided herein.

A protein of the invention, or a derivative, fragment, analog, homolog or ortholog thereof, may be utilized as an immunogen in the generation of antibodies that immunospecifically bind these protein components.

Various procedures known within the art may be used for the production of polyclonal or monoclonal antibodies directed against a protein of the invention, or against derivatives, fragments, analogs homologs or orthologs thereof (see, for example, Antibodies: A Laboratory Manual, Harlow E, and Lane D, 1988, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, incorporated herein by reference). Some of these antibodies are discussed below.

5.13.1 Polyclonal Antibodies

5

10

15

20

25

30

35 -

For the production of polyclonal antibodies, various suitable host animals (e.g., rabbit, goat, mouse or other mammal) may be immunized by one or more injections with the native protein, a synthetic variant thereof, or a derivative of the foregoing. An appropriate immunogenic preparation can contain, for example, the naturally occurring immunogenic protein, a chemically synthesized polypeptide representing the immunogenic protein, or a recombinantly expressed immunogenic protein. Furthermore, the protein may be conjugated to a second protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. The preparation can further include an adjuvant. Various adjuvants used to increase the immunological response include, but are not limited to, Freund's (complete and incomplete), mineral gels (e.g., aluminum hydroxide), surface active substances (e.g., lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions. dinitrophenol, etc.), adjuvants usable in humans such as Bacille Calmette-Guerin and Corynebacterium parvum, or similar immunostimulatory agents. Additional examples of adjuvants which can be employed include MPL-TDM adjuvant (monophosphoryl Lipid A. synthetic trehalose dicorynomycolate).

The polyclonal antibody molecules directed against the immunogenic protein can be isolated from the mammal (e.g., from the blood) and further purified by well known techniques, such as affinity chromatography using protein A or protein G, which provide primarily the IgG fraction of immune serum. Subsequently, or alternatively, the specific antigen which is the target of the immunoglobulin sought, or an epitope thereof, may be immobilized on a column to purify the immune specific antibody by immunoaffinity chromatography. Purification of immunoglobulins is discussed, for example, by D. Wilkinson (The Scientist, published by The Scientist, Inc., Philadelphia PA, Vol. 14, No. 8 (April 17, 2000), pp. 25-28).

5.13.2 Monoclonal Antibodies

5

10

15

20

25

30

35

The term "monoclonal antibody" (MAb) or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one molecular species of antibody molecule consisting of a unique light chain gene product and a unique heavy chain gene product. In particular, the complementarity determining regions (CDRs) of the monoclonal antibody are identical in all the molecules of the population. MAbs thus contain an antigen binding site capable of immunoreacting with a particular epitope of the antigen characterized by a unique binding affinity for it.

Monoclonal antibodies can be prepared using hybridoma methods, such as those described by Kohler and Milstein, <u>Nature</u>, <u>256</u>:495 (1975). In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes can be immunized in vitro.

The immunizing agent will typically include the protein antigen, a fragment thereof or a fusion protein thereof. Generally, either peripheral blood lymphocytes are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell (Goding, Monoclonal Antibodies: Principles and Practice, Academic Press, (1986) pp. 59-103). Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells can be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

Preferred immortalized cell lines are those that fuse efficiently, support stable high level expression of antibody by the selected antibody-producing cells, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma lines, which can be obtained, for instance, from the Salk Institute Cell Distribution Center, San Diego, California and the American Type Culture Collection, Manassas, Virginia. Human myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies (Kozbor, <u>J. Immunol.</u>, <u>133</u>:3001 (1984); Brodeur et al., <u>Monoclonal Antibody Production Techniques and Applications</u>, Marcel Dekker, Inc., New York, (1987) pp. 51-63).

5

10

15

20

25

30

35

The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against the antigen. Preferably, the binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an in vitro binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunoabsorbent assay (ELISA). Such techniques and assays are known in the art. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and Pollard, Anal. Biochem., 107:220 (1980). Preferably, antibodies having a high degree of specificity and a high binding affinity for the target antigen are isolated.

After the desired hybridoma cells are identified, the clones can be subcloned by limiting dilution procedures and grown by standard methods. Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640 medium.

Alternatively, the hybridoma cells can be grown in vivo as ascites in a mammal.

The monoclonal antibodies secreted by the subclones can be isolated or purified from the culture medium or ascites fluid by conventional immunoglobulin purification procedures such as, for example, protein A-Sepharose, hydroxylapatite chromatography, gel electrophoresis, dialysis, or affinity chromatography.

The monoclonal antibodies can also be made by recombinant DNA methods, such as those described in U.S. Patent No. 4,816,567. DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA can be placed into expression vectors, which are then transfected into host cells such as simian COS cells, Chinese hamster ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also can be modified, for

example, by substituting the coding sequence for human heavy and light chain constant domains in place of the homologous murine sequences (U.S. Patent No. 4,816,567; Morrison, Nature 368, 812-13 (1994)) or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. Such a non-immunoglobulin polypeptide can be substituted for the constant domains of an antibody of the invention, or can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

5.13.2 Humanized Antibodies

10 The antibodies directed against the protein antigens of the invention can further comprise humanized antibodies or human antibodies. These antibodies are suitable for administration to humans without engendering an immune response by the human against the administered immunoglobulin. Humanized forms of antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab')2 or other antigen-15 binding subsequences of antibodies) that are principally comprised of the sequence of a human immunoglobulin, and contain minimal sequence derived from a non-human immunoglobulin. Humanization can be performed following the method of Winter and co-workers (Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-327 (1988); Verhoeyen et al., Science, 239:1534-1536 (1988)), by substituting rodent CDRs or CDR sequences for the 20 corresponding sequences of a human antibody. (See also U.S. Patent No. 5,225,539.) In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies can also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable 25 domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the framework regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin (Jones et al., 1986; Riechmann et al., 1988; and Presta, Curr. Op. Struct. Bjol., 30 2:593-596 (1992)).

5.13.3 Human Antibodies

35

Fully human antibodies relate to antibody molecules in which essentially the entire sequences of both the light chain and the heavy chain, including the CDRs, arise from human genes. Such antibodies are termed "human antibodies", or "fully human antibodies" herein.

Human monoclonal antibodies can be prepared by the trioma technique; the human B-cell hybridoma technique (see Kozbor, et al., 1983 Immunol Today 4: 72) and the EBV hybridoma technique to produce human monoclonal antibodies (see Cole, et al., 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96). Human monoclonal antibodies may be utilized in the practice of the present invention and may be produced by using human hybridomas (see Cote, et al., 1983. Proc Natl Acad Sci USA 80: 2026-2030) or by transforming human B-cells with Epstein Barr Virus in vitro (see Cole, et al., 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96).

In addition, human antibodies can also be produced using additional techniques, including phage display libraries (Hoogenboom and Winter, <u>J. Mol. Biol.</u>, <u>227</u>:381 (1991); Marks et al., <u>J. Mol. Biol.</u>, <u>222</u>:581 (1991)). Similarly, human antibodies can be made by introducing human immunoglobulin loci into transgenic animals, *e.g.*, mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, which closely resembles that seen in humans in all respects, including gene rearrangement, assembly, and antibody repertoire. This approach is described, for example, in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, and in Marks et al. (<u>Bio/Technology 10</u>, 779-783 (1992)); Lonberg et al. (<u>Nature 368</u> 856-859 (1994)); Morrison (<u>Nature 368</u>, 812-13 (1994)); Fishwild et al., (<u>Nature Biotechnology 14</u>, 845-51 (1996)); Neuberger (<u>Nature Biotechnology 14</u>, 826 (1996)); and Lonberg and Huszar (<u>Intern. Rev. Immunol.</u> 13 65-93 (1995)).

Human antibodies may additionally be produced using transgenic nonhuman animals which are modified so as to produce fully human antibodies rather than the animal's endogenous antibodies in response to challenge by an antigen. (See PCT publication WO94/02602). The endogenous genes encoding the heavy and light immunoglobulin chains in the nonhuman host have been incapacitated, and active loci encoding human heavy and light chain immunoglobulins are inserted into the host's genome. The human genes are incorporated, for example, using yeast artificial chromosomes containing the requisite human DNA segments. An animal which provides all the desired modifications is then obtained as progeny by crossbreeding intermediate transgenic animals containing fewer than the full complement of the modifications. The preferred embodiment of such a nonhuman animal is a mouse, and is termed the XenomouseTM as disclosed in PCT publications WO 96/33735 and WO 96/34096. This animal produces B cells which secrete fully human immunoglobulins. The antibodies can be obtained directly from the animal after immunization with an immunogen of interest, as, for example, a preparation of a polyclonal antibody, or alternatively from immortalized B cells derived from the animal, such as hybridomas producing monoclonal antibodies. Additionally, the genes encoding the

immunoglobulins with human variable regions can be recovered and expressed to obtain the antibodies directly, or can be further modified to obtain analogs of antibodies such as, for example, single chain Fv molecules.

An example of a method of producing a nonhuman host, exemplified as a mouse, lacking expression of an endogenous immunoglobulin heavy chain is disclosed in U.S. Patent No. 5,939,598. It can be obtained by a method including deleting the J segment genes from at least one endogenous heavy chain locus in an embryonic stem cell to prevent rearrangement of the locus and to prevent formation of a transcript of a rearranged immunoglobulin heavy chain locus, the deletion being effected by a targeting vector containing a gene encoding a selectable marker; and producing from the embryonic stem cell a transgenic mouse whose somatic and germ cells contain the gene encoding the selectable marker.

A method for producing an antibody of interest, such as a human antibody, is disclosed in U.S. Patent No. 5,916,771. It includes introducing an expression vector that contains a nucleotide sequence encoding a heavy chain into one mammalian host cell in culture, introducing an expression vector containing a nucleotide sequence encoding a light chain into another mammalian host cell, and fusing the two cells to form a hybrid cell. The hybrid cell expresses an antibody containing the heavy chain and the light chain.

In a further improvement on this procedure, a method for identifying a clinically relevant epitope on an immunogen, and a correlative method for selecting an antibody that binds immunospecifically to the relevant epitope with high affinity, are disclosed in PCT publication WO 99/53049.

5.13.4 Fab Fragment's and Single Chain Antibodies

According to the invention, techniques can be adapted for the production of single-chain antibodies specific to an antigenic protein of the invention (see e.g., U.S. Patent No. 4,946,778). In addition, methods can be adapted for the construction of F_{ab} expression libraries (see e.g., Huse, et al., 1989 Science 246: 1275-1281) to allow rapid and effective identification of monoclonal F_{ab} fragments with the desired specificity for a protein or derivatives, fragments, analogs or homologs thereof. Antibody fragments that contain the idiotypes to a protein antigen may be produced by techniques known in the art including, but not limited to: (i) an $F_{(ab')2}$ fragment produced by pepsin digestion of an antibody molecule; (ii) an F_{ab} fragment generated by reducing the disulfide bridges of an $F_{(ab')2}$ fragment; (iii) an F_{ab} fragment generated by the treatment of the antibody molecule with papain and a reducing agent and (iv) F_v fragments.

5.13.5 Bispecific Antibodies

5

10

15

20

25

30

35

Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens. In the present case, one of the binding specificities is for an antigenic protein of the invention. The second binding target is any other antigen, and advantageously is a cell-surface protein or receptor or receptor subunit.

5

10

15

20

25

30

35

Methods for making bispecific antibodies are known in the art. Traditionally, the recombinant production of bispecific antibodies is based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have different specificities (Milstein and Cuello, Nature, 305:537-539 (1983)). Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of ten different antibody molecules, of which only one has the correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker et al., 1991 EMBO J., 10:3655-3659.

Antibody variable domains with the desired binding specificities (antibody-antigen combining sites) can be fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in at least one of the fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host organism. For further details of generating bispecific antibodies see, for example, Suresh et al., Methods in Enzymology, 121:210 (1986).

According to another approach described in WO 96/27011, the interface between a pair of antibody molecules can be engineered to maximize the percentage of heterodimers which are recovered from recombinant cell culture. The preferred interface comprises at least a part of the CH3 region of an antibody constant domain. In this method, one or more small amino acid side chains from the interface of the first antibody molecule are replaced with larger side chains (e.g. tyrosine or tryptophan). Compensatory "cavities" of identical or similar size to the large side chain(s) are created on the interface of the second antibody molecule by replacing large amino acid side chains with smaller ones (e.g. alanine or threonine). This provides a mechanism for increasing the yield of the heterodimer over other unwanted end-products such as homodimers.

Bispecific antibodies can be prepared as full length antibodies or antibody fragments (e.g. F(ab')₂ bispecific antibodies). Techniques for generating bispecific antibodies from antibody fragments have been described in the literature. For example, bispecific antibodies can be prepared using chemical linkage. Brennan et al., Science 229:81 (1985) describe a procedure

wherein intact antibodies are proteolytically cleaved to generate F(ab')₂ fragments. These fragments are reduced in the presence of the dithiol complexing agent sodium arsenite to stabilize vicinal dithiols and prevent intermolecular disulfide formation. The Fab' fragments generated are then converted to thionitrobenzoate (TNB) derivatives. One of the Fab'-TNB derivatives is then reconverted to the Fab'-thiol by reduction with mercaptoethylamine and is mixed with an equimolar amount of the other Fab'-TNB derivative to form the bispecific antibody. The bispecific antibodies produced can be used as agents for the selective immobilization of enzymes.

Additionally, Fab' fragments can be directly recovered from E. coli and chemically coupled to form bispecific antibodies. Shalaby et al., J. Exp. Med. 175:217-225 (1992) describe the production of a fully humanized bispecific antibody F(ab')₂ molecule. Each Fab' fragment was separately secreted from E. coli and subjected to directed chemical coupling in vitro to form the bispecific antibody. The bispecific antibody thus formed was able to bind to cells overexpressing the ErbB2 receptor and normal human T cells, as well as trigger the lytic activity of human cytotoxic lymphocytes against human breast tumor targets.

Various techniques for making and isolating bispecific antibody fragments directly from recombinant cell culture have also been described. For example, bispecific antibodies have been produced using leucine zippers. Kostelny et al., J. Immunol. 148(5):1547-1553 (1992). The leucine zipper peptides from the Fos and Jun proteins were linked to the Fab' portions of two different antibodies by gene fusion. The antibody homodimers were reduced at the hinge region to form monomers and then re-oxidized to form the antibody heterodimers. This method can also be utilized for the production of antibody homodimers. The "diabody" technology described by Hollinger et al., Proc. Natl. Acad. Sci. USA 90:6444-6448 (1993) has provided an alternative mechanism for making bispecific antibody fragments. The fragments comprise a heavy-chain variable domain (V_H) connected to a light-chain variable domain (V_L) by a linker which is too short to allow pairing between the two domains on the same chain. Accordingly, the V_H and V_L domains of one fragment are forced to pair with the complementary V_L and V_H domains of another fragment, thereby forming two antigen-binding sites. Another strategy for making bispecific antibody fragments by the use of single-chain Fv (sFv) dimers has also been reported. See, Gruber et al., J. Immunol. 152:5368 (1994).

Antibodies with more than two valencies are contemplated. For example, trispecific antibodies can be prepared. Tutt et al., <u>J. Immunol.</u> 147:60 (1991). Exemplary bispecific antibodies can bind to two different epitopes, at least one of which originates in the protein antigen of the invention. Alternatively, an anti-antigenic arm of an immunoglobulin molecule can be combined with an arm which binds to a triggering molecule on

a leukocyte such as a T-cell receptor molecule (e.g. CD2, CD3, CD28, or B7), or Fc receptors for IgG (Fc R), such as Fc RI (CD64), Fc RII (CD32) and Fc RIII (CD16) so as to focus cellular defense mechanisms to the cell expressing the particular antigen. Bispecific antibodies can also be used to direct cytotoxic agents to cells which express a particular antigen. These antibodies possess an antigen-binding arm and an arm which binds a cytotoxic agent or a radionuclide chelator, such as EOTUBE, DPTA, DOTA, or TETA. Another bispecific antibody of interest binds the protein antigen described herein and further binds tissue factor (TF).

5.13.6 Heteroconjugate Antibodies

10

15

20

25

30

35

Heteroconjugate antibodies are also within the scope of the present invention. Heteroconjugate antibodies are composed of two covalently joined antibodies. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells (U.S. Patent No. 4,676,980), and for treatment of HIV infection (WO 91/00360; WO 92/200373; EP 03089). It is contemplated that the antibodies can be prepared in vitro using known methods in synthetic protein chemistry, including those involving crosslinking agents. For example, immunotoxins can be constructed using a disulfide exchange reaction or by forming a thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

5.13.7 Effector Function Engineering

It can be desirable to modify the antibody of the invention with respect to effector function, so as to enhance, e.g., the effectiveness of the antibody in treating cancer. For example, cysteine residue(s) can be introduced into the Fc region, thereby allowing interchain disulfide bond formation in this region. The homodimeric antibody thus generated can have improved internalization capability and/or increased complement-mediated cell killing and antibody-dependent cellular cytotoxicity (ADCC). See Caron et al., J. Exp Med., 176: 1191-1195 (1992) and Shopes, J. Immunol., 148: 2918-2922 (1992). Homodimeric antibodies with enhanced antitumor activity can also be prepared using heterobifunctional cross-linkers as described in Wolff et al. Cancer Research, 53: 2560-2565 (1993). Alternatively, an antibody can be engineered that has dual Fc regions and can thereby have enhanced complement lysis and ADCC capabilities. See Stevenson et al., Anti-Cancer Drug Design, 3: 219-230 (1989).

5.13.8 Immunoconjugates

The invention also pertains to immunoconjugates comprising an antibody conjugated to a cytotoxic agent such as a chemotherapeutic agent, toxin (e.g., an enzymatically active toxin of

bacterial, fungal, plant, or animal origin, or fragments thereof), or a radioactive isotope (i.e., a radioconjugate).

Chemotherapeutic agents useful in the generation of such immunoconjugates have been described above. Enzymatically active toxins and fragments thereof that can be used include diphtheria A chain, nonbinding active fragments of diphtheria toxin, exotoxin A chain (from Pseudomonas aeruginosa), ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, Aleurites fordii proteins, dianthin proteins, Phytolaca americana proteins (PAPI, PAPII, and PAP-S), momordica charantia inhibitor, curcin, crotin, sapaonaria officinalis inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin, and the tricothecenes. A variety of radionuclides are available for the production of radioconjugated antibodies. Examples include ²¹²Bi, ¹³¹I, ¹³¹In, ⁹⁰Y, and ¹⁸⁶Re.

Conjugates of the antibody and cytotoxic agent are made using a variety of bifunctional protein-coupling agents such as N-succinimidyl-3-(2-pyridyldithiol) propionate (SPDP), iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutareldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta et al., Science, 238: 1098 (1987). Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionucleotide to the antibody. See WO94/11026.

In another embodiment, the antibody can be conjugated to a "receptor" (such streptavidin) for utilization in tumor pretargeting wherein the antibody-receptor conjugate is administered to the patient, followed by removal of unbound conjugate from the circulation using a clearing agent and then administration of a "ligand" (e.g., avidin) that is in turn conjugated to a cytotoxic agent.

4.14 COMPUTER READABLE SEQUENCES

5

10

15

20

25

30

35

In one application of this embodiment, a nucleotide sequence of the present invention can be recorded on computer readable media. As used herein, "computer readable media" refers to any medium which can be read and accessed directly by a computer. Such media include, but are not limited to: magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as CD-ROM; electrical storage media such as RAM and ROM; and hybrids of these categories such as magnetic/optical storage media. A skilled

artisan can readily appreciate how any of the presently known computer readable mediums can be used to create a manufacture comprising computer readable medium having recorded thereon a nucleotide sequence of the present invention. As used herein, "recorded" refers to a process for storing information on computer readable medium. A skilled artisan can readily adopt any of the presently known methods for recording information on computer readable medium to generate manufactures comprising the nucleotide sequence information of the present invention.

A variety of data storage structures are available to a skilled artisan for creating a computer readable medium having recorded thereon a nucleotide sequence of the present invention. The choice of the data storage structure will generally be based on the means chosen to access the stored information. In addition, a variety of data processor programs and formats can be used to store the nucleotide sequence information of the present invention on computer readable medium. The sequence information can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and Microsoft Word, or represented in the form of an ASCII file, stored in a database application, such as DB2, Sybase, Oracle, or the like. A skilled artisan can readily adapt any number of data processor structuring formats (e.g. text file or database) in order to obtain computer readable medium having recorded thereon the nucleotide sequence information of the present invention.

By providing any of the nucleotide sequences SEQ ID NO:1-1350 or a representative fragment thereof; or a nucleotide sequence at least 95% identical to any of the nucleotide sequences of SEQ ID NO:1-1350 in computer readable form, a skilled artisan can routinely access the sequence information for a variety of purposes. Computer software is publicly available which allows a skilled artisan to access sequence information provided in a computer readable medium. The examples which follow demonstrate how software which implements the BLAST (Altschul et al., J. Mol. Biol. 215:403-410 (1990)) and BLAZE (Brutlag et al., Comp. Chem. 17:203-207 (1993)) search algorithms on a Sybase system is used to identify open reading frames (ORFs) within a nucleic acid sequence. Such ORFs may be protein encoding fragments and may be useful in producing commercially important proteins such as enzymes used in fermentation reactions and in the production of commercially useful metabolites.

As used herein, "a computer-based system" refers to the hardware means, software means, and data storage means used to analyze the nucleotide sequence information of the present invention. The minimum hardware means of the computer-based systems of the present invention comprises a central processing unit (CPU), input means, output means, and data storage means. A skilled artisan can readily appreciate that any one of the currently available computer-based systems are suitable for use in the present invention. As stated above, the computer-based systems of the present invention comprise a data storage means having stored

therein a nucleotide sequence of the present invention and the necessary hardware means and software means for supporting and implementing a search means. As used herein, "data storage means" refers to memory which can store nucleotide sequence information of the present invention, or a memory access means which can access manufactures having recorded thereon the nucleotide sequence information of the present invention.

As used herein, "search means" refers to one or more programs which are implemented on the computer-based system to compare a target sequence or target structural motif with the sequence information stored within the data storage means. Search means are used to identify fragments or regions of a known sequence which match a particular target sequence or target motif. A variety of known algorithms are disclosed publicly and a variety of commercially available software for conducting search means are and can be used in the computer-based systems of the present invention. Examples of such software includes, but is not limited to. Smith-Waterman, MacPattern (EMBL), BLASTN and BLASTA (NPOLYPEPTIDEIA). A skilled artisan can readily recognize that any one of the available algorithms or implementing software packages for conducting homology searches can be adapted for use in the present computer-based systems. As used herein, a "target sequence" can be any nucleic acid or amino acid sequence of six or more nucleotides or two or more amino acids. A skilled artisan can readily recognize that the longer a target sequence is, the less likely a target sequence will be present as a random occurrence in the database. The most preferred sequence length of a target sequence is from about 10 to 300 amino acids, more preferably from about 30 to 100 nucleotide residues. However, it is well recognized that searches for commercially important fragments, such as sequence fragments involved in gene expression and protein processing, may be of shorter length.

As used herein, "a target structural motif," or "target motif," refers to any rationally selected sequence or combination of sequences in which the sequence(s) are chosen based on a three-dimensional configuration which is formed upon the folding of the target motif. There are a variety of target motifs known in the art. Protein target motifs include, but are not limited to, enzyme active sites and signal sequences. Nucleic acid target motifs include, but are not limited to, promoter sequences, hairpin structures and inducible expression elements (protein binding sequences).

4.15 TRIPLE HELIX FORMATION

5

10

15

20

25

30

35

In addition, the fragments of the present invention, as broadly described, can be used to control gene expression through triple helix formation or antisense DNA or RNA, both of which methods are based on the binding of a polynucleotide sequence to DNA or RNA.

Polynucleotides suitable for use in these methods are preferably 20 to 40 bases in length and are designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 15241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Olmno, J. Neurochem. 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix-formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems. Information contained in the sequences of the present invention is necessary for the design of an antisense or triple helix oligonucleotide.

4.16 DIAGNOSTIC ASSAYS AND KITS

5

10

15

20

25

30

35

The present invention further provides methods to identify the presence or expression of one of the ORFs of the present invention, or homolog thereof, in a test sample, using a nucleic acid probe or antibodies of the present invention, optionally conjugated or otherwise associated with a suitable label.

In general, methods for detecting a polynucleotide of the invention can comprise contacting a sample with a compound that binds to and forms a complex with the polynucleotide for a period sufficient to form the complex, and detecting the complex, so that if a complex is detected, a polynucleotide of the invention is detected in the sample. Such methods can also comprise contacting a sample under stringent hybridization conditions with nucleic acid primers that anneal to a polynucleotide of the invention under such conditions, and amplifying annealed polynucleotides, so that if a polynucleotide is amplified, a polynucleotide of the invention is detected in the sample.

In general, methods for detecting a polypeptide of the invention can comprise contacting a sample with a compound that binds to and forms a complex with the polypeptide for a period sufficient to form the complex, and detecting the complex, so that if a complex is detected, a polypeptide of the invention is detected in the sample.

In detail, such methods comprise incubating a test sample with one or more of the antibodies or one or more of the nucleic acid probes of the present invention and assaying for binding of the nucleic acid probes or antibodies to components within the test sample.

Conditions for incubating a nucleic acid probe or antibody with a test sample vary.

Incubation conditions depend on the format employed in the assay, the detection methods employed, and the type and nature of the nucleic acid probe or antibody used in the assay. One skilled in the art will recognize that any one of the commonly available hybridization,

amplification or immunological assay formats can readily be adapted to employ the nucleic acid probes or antibodies of the present invention. Examples of such assays can be found in Chard, T., An Introduction to Radioimmunoassay and Related Techniques, Elsevier Science Publishers, Amsterdam, The Netherlands (1986); Bullock, G.R. et al., Techniques in Immunocytochemistry, Academic Press, Orlando, FL Vol. 1 (1982), Vol. 2 (1983), Vol. 3 (1985); Tijssen, P., Practice and Theory of immunoassays: Laboratory Techniques in Biochemistry and Molecular Biology, Elsevier Science Publishers, Amsterdam, The Netherlands (1985). The test samples of the present invention include cells, protein or membrane extracts of cells, or biological fluids such as sputum, blood, serum, plasma, or urine. The test sample used in the above-described method will vary based on the assay format, nature of the detection method and the tissues, cells or extracts used as the sample to be assayed. Methods for preparing protein extracts or membrane extracts of cells are well known in the art and can be readily be adapted in order to obtain a sample which is compatible with the system utilized.

In another embodiment of the present invention, kits are provided which contain the necessary reagents to carry out the assays of the present invention. Specifically, the invention provides a compartment kit to receive, in close confinement, one or more containers which comprises: (a) a first container comprising one of the probes or antibodies of the present invention; and (b) one or more other containers comprising one or more of the following: wash reagents, reagents capable of detecting presence of a bound probe or antibody.

In detail, a compartment kit includes any kit in which reagents are contained in separate containers. Such containers include small glass containers, plastic containers or strips of plastic or paper. Such containers allows one to efficiently transfer reagents from one compartment to another compartment such that the samples and reagents are not cross-contaminated, and the agents or solutions of each container can be added in a quantitative fashion from one compartment to another. Such containers will include a container which will accept the test sample, a container which contains the antibodies used in the assay, containers which contain wash reagents (such as phosphate buffered saline, Tris-buffers, etc.), and containers which contain the reagents used to detect the bound antibody or probe. Types of detection reagents include labeled nucleic acid probes, labeled secondary antibodies, or in the alternative, if the primary antibody is labeled, the enzymatic, or antibody binding reagents which are capable of reacting with the labeled antibody. One skilled in the art will readily recognize that the disclosed probes and antibodies of the present invention can be readily incorporated into one of the established kit formats which are well known in the art.

4.17 MEDICAL IMAGING

The novel polypeptides and binding partners of the invention are useful in medical imaging of sites expressing the molecules of the invention (e.g., where the polypeptide of the invention is involved in the immune response, for imaging sites of inflammation or infection). See, e.g., Kunkel et al., U.S. Pat. NO. 5,413,778. Such methods involve chemical attachment of a labeling or imaging agent, administration of the labeled polypeptide to a subject in a pharmaceutically acceptable carrier, and imaging the labeled polypeptide in vivo at the target site.

4.18 SCREENING ASSAYS

(b)

5

10

15

20

25

30

35

Using the isolated proteins and polynucleotides of the invention, the present invention further provides methods of obtaining and identifying agents which bind to a polypeptide encoded by an ORF corresponding to any of the nucleotide sequences set forth in SEQ ID NO:1-1350, or bind to a specific domain of the polypeptide encoded by the nucleic acid. In detail, said method comprises the steps of:

- (a) contacting an agent with an isolated protein encoded by an ORF of the present invention, or nucleic acid of the invention; and
- determining whether the agent binds to said protein or said nucleic acid. In general, therefore, such methods for identifying compounds that bind to a polynucleotide of the invention can comprise contacting a compound with a polynucleotide of the invention for a time sufficient to form a polynucleotide/compound complex, and detecting

the complex, so that if a polynucleotide/compound complex is detected, a compound that binds to a polynucleotide of the invention is identified.

Likewise, in general, therefore, such methods for identifying compounds that bind to a polypeptide of the invention can comprise contacting a compound with a polypeptide of the invention for a time sufficient to form a polypeptide/compound complex, and detecting the complex, so that if a polypeptide/compound complex is detected, a compound that binds to a polynucleotide of the invention is identified.

Methods for identifying compounds that bind to a polypeptide of the invention can also comprise contacting a compound with a polypeptide of the invention in a cell for a time sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a receptor gene sequence in the cell, and detecting the complex by detecting reporter gene sequence expression, so that if a polypeptide/compound complex is detected, a compound that binds a polypeptide of the invention is identified.

Compounds identified via such methods can include compounds which modulate the activity of a polypeptide of the invention (that is, increase or decrease its activity, relative to

activity observed in the absence of the compound). Alternatively, compounds identified via such methods can include compounds which modulate the expression of a polynucleotide of the invention (that is, increase or decrease expression relative to expression levels observed in the absence of the compound). Compounds, such as compounds identified via the methods of the invention, can be tested using standard assays well known to those of skill in the art for their ability to modulate activity/expression.

5

10

15

20

25

30

35

The agents screened in the above assay can be, but are not limited to, peptides, carbohydrates, vitamin derivatives, or other pharmaceutical agents. The agents can be selected and screened at random or rationally selected or designed using protein modeling techniques.

For random screening, agents such as peptides, carbohydrates, pharmaceutical agents and the like are selected at random and are assayed for their ability to bind to the protein encoded by the ORF of the present invention. Alternatively, agents may be rationally selected or designed. As used herein, an agent is said to be "rationally selected or designed" when the agent is chosen based on the configuration of the particular protein. For example, one skilled in the art can readily adapt currently available procedures to generate peptides, pharmaceutical agents and the like, capable of binding to a specific peptide sequence, in order to generate rationally designed antipeptide peptides, for example see Hurby et al., Application of Synthetic Peptides: Antisense Peptides," In Synthetic Peptides, A User's Guide, W.H. Freeman, NY (1992), pp. 289-307, and Kaspczak et al., Biochemistry 28:9230-8 (1989), or pharmaceutical agents, or the like.

In addition to the foregoing, one class of agents of the present invention, as broadly described, can be used to control gene expression through binding to one of the ORFs or EMFs of the present invention. As described above, such agents can be randomly screened or rationally designed/selected. Targeting the ORF or EMF allows a skilled artisan to design sequence specific or element specific agents, modulating the expression of either a single ORF or multiple ORFs which rely on the same EMF for expression control. One class of DNA binding agents are agents which contain base residues which hybridize or form a triple helix formation by binding to DNA or RNA. Such agents can be based on the classic phosphodiester, ribonucleic acid backbone, or can be a variety of sulfhydryl or polymeric derivatives which have base attachment capacity.

Agents suitable for use in these methods preferably contain 20 to 40 bases and are designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Okano, J. Neurochem. 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix-formation optimally results in a shut-off of RNA transcription

from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems. Information contained in the sequences of the present invention is necessary for the design of an antisense or triple helix oligonucleotide and other DNA binding agents.

Agents which bind to a protein encoded by one of the ORFs of the present invention can be used as a diagnostic agent. Agents which bind to a protein encoded by one of the ORFs of the present invention can be formulated using known techniques to generate a pharmaceutical composition.

4.19 USE OF NUCLEIC ACIDS AS PROBES

5

10

15

20

25

30

Another aspect of the subject invention is to provide for polypeptide-specific nucleic acid hybridization probes capable of hybridizing with naturally occurring nucleotide sequences. The hybridization probes of the subject invention may be derived from any of the nucleotide sequences SEQ ID NO:1-1350. Because the corresponding gene is only expressed in a limited number of tissues, a hybridization probe derived from of any of the nucleotide sequences SEQ ID NO:1-1350 can be used as an indicator of the presence of RNA of cell type of such a tissue in a sample.

Any suitable hybridization technique can be employed, such as, for example, in situ hybridization. PCR as described in US Patents Nos. 4,683,195 and 4,965,188 provides additional uses for oligonucleotides based upon the nucleotide sequences. Such probes used in PCR may be of recombinant origin, may be chemically synthesized, or a mixture of both. The probe will comprise a discrete nucleotide sequence for the detection of identical sequences or a degenerate pool of possible sequences for identification of closely related genomic sequences.

Other means for producing specific hybridization probes for nucleic acids include the cloning of nucleic acid sequences into vectors for the production of mRNA probes. Such vectors are known in the art and are commercially available and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerase as T7 or SP6 RNA polymerase and the appropriate radioactively labeled nucleotides. The nucleotide sequences may be used to construct hybridization probes for mapping their respective genomic sequences. The nucleotide sequence provided herein may be mapped to a chromosome or specific regions of a chromosome using well known genetic and/or chromosomal mapping techniques. These techniques include in situ hybridization, linkage analysis against known chromosomal markers, hybridization screening with libraries or flow-sorted chromosomal preparations specific to known chromosomes, and the like. The technique of fluorescent in situ hybridization of

chromosome spreads has been described, among other places, in Verma et al (1988) Human Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York NY.

5

10

15

20

25

30

Fluorescent in situ hybridization of chromosomal preparations and other physical chromosome mapping techniques may be correlated with additional genetic map data. Examples of genetic map data can be found in the 1994 Genome Issue of Science (265:1981f). Correlation between the location of a nucleic acid on a physical chromosomal map and a specific disease (or predisposition to a specific disease) may help delimit the region of DNA associated with that genetic disease. The nucleotide sequences of the subject invention may be used to detect differences in gene sequences between normal, carrier or affected individuals.

4.20 PREPARATION OF SUPPORT BOUND OLIGONUCLEOTIDES

Oligonucleotides, *i.e.*, small nucleic acid segments, may be readily prepared by, for example, directly synthesizing the oligonucleotide by chemical means, as is commonly practiced using an automated oligonucleotide synthesizer.

Support bound oligonucleotides may be prepared by any of the methods known to those of skill in the art using any suitable support such as glass, polystyrene or Teflon. One strategy is to precisely spot oligonucleotides synthesized by standard synthesizers. Immobilization can be achieved using passive adsorption (Inouye & Hondo, (1990) J. Clin. Microbiol. 28(6) 1469-72); using UV light (Nagata *et al.*, 1985; Dahlen *et al.*, 1987; Morrissey & Collins, (1989) Mol. Cell Probes 3(2) 189-207) or by covalent binding of base modified DNA (Keller *et al.*, 1988; 1989); all references being specifically incorporated herein.

Another strategy that may be employed is the use of the strong biotin-streptavidin interaction as a linker. For example, Broude *et al.* (1994) Proc. Natl. Acad. Sci. USA 91(8) 3072-6, describe the use of biotinylated probes, although these are duplex probes, that are immobilized on streptavidin-coated magnetic beads. Streptavidin-coated beads may be purchased from Dynal, Oslo. Of course, this same linking chemistry is applicable to coating any surface with streptavidin. Biotinylated probes may be purchased from various sources, such as, *e.g.*, Operon Technologies (Alameda, CA).

Nunc Laboratories (Naperville, IL) is also selling suitable material that could be used. Nunc Laboratories have developed a method by which DNA can be covalently bound to the microwell surface termed Covalink NH. CovaLink NH is a polystyrene surface grafted with secondary amino groups (>NH) that serve as bridge-heads for further covalent coupling. CovaLink Modules may be purchased from Nunc Laboratories. DNA molecules may be bound to CovaLink exclusively at the 5'-end by a phosphoramidate bond, allowing immobilization of more than 1 pmol of DNA (Rasmussen et al., (1991) Anal. Biochem. 198(1) 138-42).

The use of CovaLink NH strips for covalent binding of DNA molecules at the 5'-end has been described (Rasmussen et al., (1991). In this technology, a phosphoramidate bond is employed (Chu et al., (1983) Nucleic Acids Res. 11(8) 6513-29). This is beneficial as immobilization using only a single covalent bond is preferred. The phosphoramidate bond joins the DNA to the CovaLink NH secondary amino groups that are positioned at the end of spacer arms covalently grafted onto the polystyrene surface through a 2 nm long spacer arm. To link an oligonucleotide to CovaLink NH via an phosphoramidate bond, the oligonucleotide terminus must have a 5'-end phosphate group. It is, perhaps, even possible for biotin to be covalently bound to CovaLink and then streptavidin used to bind the probes.

5

10

15

20

25

30

More specifically, the linkage method includes dissolving DNA in water (7.5 ng/ul) and denaturing for 10 min. at 95°C and cooling on ice for 10 min. Ice-cold 0.1 M 1-methylimidazole, pH 7.0 (1-MeIm₇), is then added to a final concentration of 10 mM 1-MeIm₇. A ss DNA solution is then dispensed into CovaLink NH strips (75 ul/well) standing on ice.

Carbodiimide 0.2 M 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDC), dissolved in 10 mM 1-MeIm₇, is made fresh and 25 ul added per well. The strips are incubated for 5 hours at 50°C. After incubation the strips are washed using, e.g., Nunc-Immuno Wash; first the wells are washed 3 times, then they are soaked with washing solution for 5 min., and finally they are washed 3 times (where in the washing solution is 0.4 N NaOH, 0.25% SDS heated to 50°C).

It is contemplated that a further suitable method for use with the present invention is that described in PCT Patent Application WO 90/03382 (Southern & Maskos), incorporated herein by reference. This method of preparing an oligonucleotide bound to a support involves attaching a nucleoside 3'-reagent through the phosphate group by a covalent phosphodiester link to aliphatic hydroxyl groups carried by the support. The oligonucleotide is then synthesized on the supported nucleoside and protecting groups removed from the synthetic oligonucleotide chain under standard conditions that do not cleave the oligonucleotide from the support. Suitable reagents include nucleoside phosphoramidite and nucleoside hydrogen phosphorate.

An on-chip strategy for the preparation of DNA probe for the preparation of DNA probe arrays may be employed. For example, addressable laser-activated photodeprotection may be employed in the chemical synthesis of oligonucleotides directly on a glass surface, as described by Fodor *et al.* (1991) Science 251(4995) 767-73, incorporated herein by reference. Probes may also be immobilized on nylon supports as described by Van Ness *et al.* (1991) Nucleic Acids Res. 19(12) 3345-50; or linked to Teflon using the method of Duncan & Cavalier (1988) Anal. Biochem. 169(1) 104-8; all references being specifically incorporated herein.

To link an oligonucleotide to a nylon support, as described by Van Ness *et al.* (1991), requires activation of the nylon surface via alkylation and selective activation of the 5'-amine of oligonucleotides with cyanuric chloride.

One particular way to prepare support bound oligonucleotides is to utilize the light-generated synthesis described by Pease *et al.*, (1994) PNAS USA 91(11) 5022-6, incorporated herein by reference). These authors used current photolithographic techniques to generate arrays of immobilized oligonucleotide probes (DNA chips). These methods, in which light is used to direct the synthesis of oligonucleotide probes in high-density, miniaturized arrays, utilize photolabile 5'-protected *N*-acyl-deoxynucleoside phosphoramidites, surface linker chemistry and versatile combinatorial synthesis strategies. A matrix of 256 spatially defined oligonucleotide probes may be generated in this manner.

4.21 PREPARATION OF NUCLEIC ACID FRAGMENTS

5

10

15

20

25

30

The nucleic acids may be obtained from any appropriate source, such as cDNAs, genomic DNA, chromosomal DNA, microdissected chromosome bands, cosmid or YAC inserts, and RNA, including mRNA without any amplification steps. For example, Sambrook *et al.* (1989) describes three protocols for the isolation of high molecular weight DNA from mammalian cells (p. 9.14-9.23).

DNA fragments may be prepared as clones in M13, plasmid or lambda vectors and/or prepared directly from genomic DNA or cDNA by PCR or other amplification methods. Samples may be prepared or dispensed in multiwell plates. About 100-1000 ng of DNA samples may be prepared in 2-500 ml of final volume.

The nucleic acids would then be fragmented by any of the methods known to those of skill in the art including, for example, using restriction enzymes as described at 9.24-9.28 of Sambrook *et al.* (1989), shearing by ultrasound and NaOH treatment.

Low pressure shearing is also appropriate, as described by Schriefer *et al.* (1990) Nucleic Acids Res. 18(24) 7455-6, incorporated herein by reference). In this method, DNA samples are passed through a small French pressure cell at a variety of low to intermediate pressures. A lever device allows controlled application of low to intermediate pressures to the cell. The results of these studies indicate that low-pressure shearing is a useful alternative to sonic and enzymatic DNA fragmentation methods.

One particularly suitable way for fragmenting DNA is contemplated to be that using the two base recognition endonuclease, CviJI, described by Fitzgerald et al. (1992) Nucleic Acids Res. 20(14) 3753-62. These authors described an approach for the rapid fragmentation and fractionation

of DNA into particular sizes that they contemplated to be suitable for shotgun cloning and sequencing.

The restriction endonuclease CviII normally cleaves the recognition sequence PuGCPy between the G and C to leave blunt ends. Atypical reaction conditions, which alter the specificity of this enzyme ($CviII^{**}$), yield a quasi-random distribution of DNA fragments form the small molecule pUC19 (2688 base pairs). Fitzgerald $et\ al.$ (1992) quantitatively evaluated the randomness of this fragmentation strategy, using a $CviII^{**}$ digest of pUC19 that was size fractionated by a rapid gel filtration method and directly ligated, without end repair, to a lac Z minus M13 cloning vector. Sequence analysis of 76 clones showed that $CviII^{**}$ restricts pyGCPy and PuGCPu, in addition to PuGCPy sites, and that new sequence data is accumulated at a rate consistent with random fragmentation.

As reported in the literature, advantages of this approach compared to sonication and agarose gel fractionation include: smaller amounts of DNA are required (0.2-0.5 ug instead of 2-5 ug); and fewer steps are involved (no preligation, end repair, chemical extraction, or agarose gel electrophoresis and elution are needed

Irrespective of the manner in which the nucleic acid fragments are obtained or prepared, it is important to denature the DNA to give single stranded pieces available for hybridization. This is achieved by incubating the DNA solution for 2-5 minutes at 80-90°C. The solution is then cooled quickly to 2°C to prevent renaturation of the DNA fragments before they are contacted with the chip. Phosphate groups must also be removed from genomic DNA by methods known in the art.

4.22 PREPARATION OF DNA ARRAYS

5

10

15

20

25

30

Arrays may be prepared by spotting DNA samples on a support such as a nylon membrane. Spotting may be performed by using arrays of metal pins (the positions of which correspond to an array of wells in a microtiter plate) to repeated by transfer of about 20 nl of a DNA solution to a nylon membrane. By offset printing, a density of dots higher than the density of the wells is achieved. One to 25 dots may be accommodated in 1 mm², depending on the type of label used. By avoiding spotting in some preselected number of rows and columns, separate subsets (subarrays) may be formed. Samples in one subarray may be the same genomic segment of DNA (or the same gene) from different individuals, or may be different, overlapped genomic clones. Each of the subarrays may represent replica spotting of the same samples. In one example, a selected gene segment may be amplified from 64 patients. For each patient, the amplified gene segment may be in one 96-well plate (all 96 wells containing the same sample). A plate for each of the 64 patients is prepared. By using a 96-pin device, all samples may be spotted on one 8 x 12 cm membrane.

Subarrays may contain 64 samples, one from each patient. Where the 96 subarrays are identical, the dot span may be 1 mm² and there may be a 1 mm space between subarrays.

Another approach is to use membranes or plates (available from NUNC, Naperville, Illinois) which may be partitioned by physical spacers *e.g.* a plastic grid molded over the membrane, the grid being similar to the sort of membrane applied to the bottom of multiwell plates, or hydrophobic strips. A fixed physical spacer is not preferred for imaging by exposure to flat phosphor-storage screens or x-ray films.

The present invention is illustrated in the following examples. Upon consideration of the present disclosure, one of skill in the art will appreciate that many other embodiments and variations may be made in the scope of the present invention. Accordingly, it is intended that the broader aspects of the present invention not be limited to the disclosure of the following examples. The present invention is not to be limited in scope by the exemplified embodiments which are intended as illustrations of single aspects of the invention, and compositions and methods which are functionally equivalent are within the scope of the invention. Indeed, numerous modifications and variations in the practice of the invention are expected to occur to those skilled in the art upon consideration of the present preferred embodiments. Consequently, the only limitations which should be placed upon the scope of the invention are those which appear in the appended claims.

All references cited within the body of the instant specification are hereby incorporated by reference in their entirety.

5.0 EXAMPLES

5

10

15

20

25

30

5.1 EXAMPLE 1

Novel Nucleic Acid Sequences Obtained From Various Libraries

A plurality of novel nucleic acids were obtained from cDNA libraries prepared from various human tissues and in some cases isolated from a genomic library derived from human chromosome using standard PCR, SBH sequence signature analysis and Sanger sequencing techniques. The inserts of the library were amplified with PCR using primers specific for the vector sequences which flank the inserts. Clones from cDNA libraries were spotted on nylon membrane filters and screened with oligonucleotide probes (e.g., 7-mers) to obtain signature sequences. The clones were clustered into groups of similar or identical sequences. Representative clones were selected for sequencing.

In some cases, the 5' sequence of the amplified inserts was then deduced using a typical Sanger sequencing protocol. PCR products were purified and subjected to fluorescent dye terminator cycle sequencing. Single pass gel sequencing was done using a 377 Applied Biosystems

(ABI) sequencer to obtain the novel nucleic acid sequences. In some cases RACE (Random Amplification of cDNA Ends) was performed to further extend the sequence in the 5' direction.

5.2 EXAMPLE 2

Novel Contigs

5

10

15

20

25

30

The novel contigs of the invention were assembled from sequences that were obtained from a cDNA library by methods described in Example 1 above, and in some cases sequences obtained from one or more public databases. The sequences for the resulting nucleic acid contigs are designated as SEQ ID NO: 1-1350 and are provided in the attached Sequence Listing. The contigs were assembled using an EST sequence as a seed. Then a recursive algorithm was used to extend the seed EST into an extended assemblage, by pulling additional sequences from different databases (i.e., Hyseq's database containing EST sequences, dbEST version 114, gb pri 114, and UniGene version 101) that belong to this assemblage. The algorithm terminated when there was no additional sequences from the above databases that would extend the assemblage. Inclusion of component sequences into the assemblage was based on a BLASTN hit to the extending assemblage with BLAST score greater than 300 and percent identity greater than 95%.

Table 3 sets forth the novel predicted polypeptides (including proteins) encoded by the novel polynucleotides (SEQ ID NO:189-282) of the present invention, and their corresponding nucleotide locations to each of SEQ ID NO: 189-282. Table 3 also indicates the method by which the polypeptide was predicted. Method A refers to a polypeptide obtained by using a software program called FASTY (available from http://fasta.bioch.virginia.edu) which selects a polypeptide based on a comparison of the translated novel polynucleotide to known polynucleotides (W.R. Pearson, Methods in Enzymology, 183:63-98 (1990), herein incorporated by reference). Method B refers to a polypeptide obtained by using a software program called GenScan for human/vertebrate sequences (available from Stanford University, Office of Technology Licensing) that predicts the polypeptide based on a probabilistic model of gene structure/compositional properties (C. Burge and S. Karlin, J. Mol. Biol., 268:78-94 (1997), incorporated herein by reference). Method C refers to a polypeptide obtained by using a Hyseq proprietary software program that translates the novel polynucleotide and its complementary strand into six possible amino acid sequences (forward and reverse frames) and chooses the polypeptide with the longest open reading frame.

The nearest neighbor results for SEQ ID NO: 1-1350 were obtained by a BLASTP version 2.0al 19MP-WashU search against Genpept release 120 and Geneseq database October 12, 2000, update 21 (Derwent), using BLAST algorithm. The nearest neighbor result showed the

closest homologue for SEQ ID NO:1-1350. The nearest neighbor results for SEQ ID NO: 1-1350 are shown in Table 2 below.

Tables 1, 2 and 3 follow. Table 1 shows the various tissue sources of SEQ ID NO: 1-1350. Table 2 shows the nearest neighbor result for the assembled contig. The nearest neighbor result shows the closest homolog with an identifiable function for each assemblage. Table 3 contains the start and stop nucleotides for the translated amino acid sequence for which each assemblage encodes. Table 3 also provides a correlation between the amino acid sequences set forth in the Sequence Listing, the nucleotide sequences set forth in the Sequence Listing and the SEQ ID NO. in USSN 09/496,914.

5

TABLE 1

Tissue Origin	RNA Source	Hyseq Library Nam	e SEQ ID NOS:
adult brain	GIBCO	AB3001	111 151 188 215 662-665 877 910 927
			976 1233 1319
adult brain	GIBCO	ABD003	41 49 74 101 111 120 132 141-142 151
			217 225 238 271 317 404 446 469 503
			513-514 535 550 564 573 666-669 798
			898 910 927 976 1067 1083 1085 1178
adult brain	Clarita	10001	1254
auun bram	Clontech	ABR001	39 216 238 327 356 535 927 1056 1121
adult brain	Clontech	ABR006	1178-1180 1199 1251
adult brain	Clontech	ABR008	74 611 949 1034 1136 14 32 41 61 81 86 89 120 132 138 145
	Cionicon	. ADKOO	147 188 197 208 225 227-239 250 300-
		1	303 312 316 328-331 340 357-362 374
		•	380 384-391 408 414 446 448 464-467
			483 488 495-496 505 512 521 535 550
			566 571 577 585 590 594 598 634 641
			658 666 683 725 742 764 767 786 801
			805 810 823 826 829 831 836 841 887-
			923 927 934 943 950-951 963 976 995
			1000-1001 1006 1026 1034 1048 1057-
			1067 1086 1088 1090 1118 1120 1122-
		j	1128 1142 1162 1181-1192 1199 1204 1218-1219 1225 1232 1253 1267 1271-
			1306 1342 1347 1349-1350
adult brain	Clontech	ABR011	49 238 1219
adult brain	BioChain	ABR012	74 238
adult brain	Invitrogen	ABR013	868 1268
adult brain	Invitrogen	ABT004	49 117 138 191 217 252 291 305 535
	1		566 596 663 670 746 798 816-819 876
			892 898 922 943 963 1034-1036 1121
cultured	Strategene	ADP001	41 74 101 138 211 238 304 537 582
preadipocytes adrenal gland	<u> </u>	ADDOO	740 798 883 943 976 1067
adrenai giand	Clontech	ADR002	49 74 101 111 120 127 151 215 238
			240-247 316 330 363-364 404 414 534-
			535 833 924-940 950 963 976 1001 1003 1067-1070 1118 1156 1193-1200
			1325
adult heart	GIBCO	AHR001	38 49 71-72 74-77 79 92 99 101 111
		1	118 129 132 138 151 158-163 182 195-
			203 215 217 238 264 269 353 384 398
			408 434-439 446 504 512-513 519 537
			562-573 577 611-614 616-619 658 661
			671-672 722 734 757-773 815 828-835
	1		874 891 898 919 926-927 976 988
	-		1021 1037 1041 1062 1067 1071 1080
			1083 1093 1122 1131 1185 1201 1254
adult kidney	GIBCO	AKD001	1308 1331 1335 41 49 51 71-74 78-85 94 100-101 103-
addit Ridiley	Gibeo	AKDOOI	107 111 119-120 138 151 157 215 217-
			218 238 250 264 294 304 384 404 440
			446 454 477 504-505 509 514 518-519
		1	535 537 564 574-583 620-627 639 653
			673-675 705 753 789 831 844 851 859
			877 909 918 927 956 963 976 1067
			1074 1083 1095 1178 1302 1331 1335
adult kidney	Invitrogen	AKT002	11-12 41 49 111-112 215-217 294 316
	1	1 ,	446 487 564 575 844 868 910 927 976
adult lung	CIRCO	AL COOL	1116
adult lung	GIBCO	ALG001	8 101 111 151 187 402 446 490 514

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
		 	518 537 545 549 580 582 592 594 634
		ĺ	640 651-652 676-678 725 851 873 918
			952 976 1042 1067 1076 1083 1152
lymph node	Clontech	ALN001	8 111 121 151 180-182 188 215 537
]		545 549 651 679-682 789 804-810 868
			873 927 952 976 1042 1059 1335
young liver	GIBCO	ALV001	8 64 79 111 186 215-216 238 446 514
	}		519 537 564 653 683-684 698 753 798
			813 833 840 858 927 976 1038-1039
			1051 1085 1224 1245 1256
adult liver	Invitrogen	ALV002	40 71 292-293 305 384 468-469 496
			505 657 675 714 753 832 844 941-942
			976 1040 1076 1256 1293
adult liver	Clontech	ALV003	976
adult ovary	Invitrogen	AOV001	8 32 36 38 41 49 51 71 74 79-80 101
			104 111 120 122-125 138 140 143-149
			151 188-190 207-212 215-217 238 264
		[316 384 409 440 445-446 496 504 512
			514 518-519 535 537 549-550 564 566
			571 580 582 600 618 638 657 667 681
			685-697 699 705 722 735-744 761 771 815 833 842-865 868 875-876 918 926-
			927 950 952 963 976 1023 1042 1048
			1051 1059 1072 1076 1083 1117 1120
	İ	}	1124 1131 1144 1174 1224 1268 1331
	ļ.		1335
adult placenta	Clontech	APL001	102 217 238 537 641 700
placenta	Invitrogen	APL002	663 851 1048
adult spleen	GIBCO	ASP001	8 45 74 111 132 140 151 185 217 238
addit spicen	dibeo	A51 001	294 414 446 477 504 514 534 545 549
		1	592 722 873 883 952 976 1041-1042
			1083 1093-1094 1152 1224
testis	GIBCO	ATS001	72 107 111 113 126 140 151 183 215
			238 446 497 537 642 701-706 811 877
			927 962 976 1083 1117 1131
adult bladder	Invitrogen	BLD001	41 151 191 402-405 409 414 496 545
			592 607 706 873 952 1178 1329-1335
bone marrow	Clontech	BMD001	8 58-62 65-68 74 79 108 111 116 137
			147 151 164-174 213-215 238 305-307
			374 404 446 460 466 516 519 534 538-
	. }		541 544-546 549-554 566 584 586 592
			596 607 610 628-629 643-645 652 707-
)		708 774-789 844 866-871 873 919 927
			952 963 976 998 1034 1042 1064 1083
			1085 1120 1132 1152 1225 1229 1268
			1307 1310
bone marrow	Clontech	BMD002	6 8 37-38 52 74 77 105 111 129 132
	ļ		210 317 510-511 545 549 581 598 628
			638 724 766 789 844 860 868 873 919
	ļ		927 952 963 968 976 1042 1111 1141
			1160-1161 1229 1266 1346
bone marrow	Clontech	BMD004	111 238 282 549 1083
adult colon	Invitrogen	CLN001	52 260 264 299 494 536 545 564 592
			844 873 877 952 976 1042 1152 1268
	<u> </u>	<u> </u>	1336-1337
adult cervix	BioChain	CVX001	49 51 129 132 151 205 207 238 332-
			335 365-367 392-401 440 466 470-471
		1	518 537 597 629 832 877 927 976 1006
		J	1085 1117 1129-1134 1192 1202-1205
dianhraam	BioChain	DIA002	1219 1309-1328 74 976 1083
diaphragm		ן אועען	147/0 1003

Tissue Origin	RNA Source	Hyseq Library Name	Tero in Mos.
endothelial cells	Strategene	EDT001	SEQ ID NOS: 32 40-41 49 74 79 101 111 120 132
ondomental cons	omatogene	LD1001	138 151 204-206 215-217 238 269 316
		[414 433 505 510 513 550 555 580 582
		1	596 675 722 745 798 814 836-841 851
j	1		918 976 1041 1043 1073 1083 1131
	1		1331
Genomic clones	Genomic DNA	EPM001	525-532 927
from the short arm	from Genetic		323 332 72 7
of chromosome 8	Research		
Genomic clones	Genomic DNA	EPM003	47 525
from the short arm	from Genetic		
of chromosome 8	Research	ì	1
Genomic clones	Genomic DNA	EPM004	525 927
from the short arm	from Genetic		
of chromosome 8	Research		
Genomic clones	Genomic DNA	EPM005	531
from the short arm	from Genetic	ĺ	İ
of chromosome 8	Research	<u></u>	1
esophagus	BioChain	ESO002	74 138 238
fetal brain	Clontech	FBR001	441-442 927
fetal brain	Clontech	FBR004	215 893 927 1001
fetal brain	Clontech	FBR006	48 61 101 120 132 138 140 147 208
			225 271 317 319 336 359 368 405-414
			519 550 571 594 686 715 722 764 824
			829 836 859 909 927 943 947 963 1057
			1067-1068 1104 1135-1140 1162 1206-
			1207 1235 1268 1288 1307-1308 1319
fetal brain	Classia	TDD 00	1338-1350
fetal brain	Clontech	FBRs03	111 446
letai trani	Invitrogen	FBT002	41 51 120 151 192-194 264 504 512
	}		535 683 761 798 820-827 844 876 909 963 976 1026 1048 1083 1144 1302
fetal heart	Invitrogen	FHR001	446 566 761
fetal kidney	Clontech	FKD001	51 74 111 127 140 151 184 294 537
	Ciomon	TREOUT	550 630-631 1319
fetal kidney	Clontech	FKD002	111 976 1083
fetal kidney	Invitrogen	FKD007	238 974
fetal lung	Clontech	FLG001	463 566 976 1074 1083 1093
fetal lung	Invitrogen	FLG003	41 238 330 407 415-416 537 573 844
			859 1048 1083 1116 1192
fetal liver-spleen	Columbia	FLS001	8 14 34-35 37 41 43 49 51 54-56 63-64
	University		69-71 74 77 79 87-90 101 107 110-111
			114 120 128-131 138 140 147 150-155
	j (197 210 215 217 225 238 312 367 384
			414 440 446 460 468 483 496 504-507
			511-515 518-519 523 533-535 537 541
	1		544-545 547-550 555-560 564 566 571
			577 582 585-586 598 636 646-647 649
			652 664 698 709-710 714 722-723 731
)	,	735-736 746-753 761 784 798 823 829
			832 844 851 858-859 868 873 876 898
			927 943 949 952 963 976 984 1002
	, !		1021 1023 1040 1042 1044 1050 1083
			1093 1116 1120 1129 1131 1144 1174
	1		1319
fetal liver-spleen	Columbia	FLS002	8 36-37 41-46 49 54 64 71 74 79 101
and an or opioon	University	1 140002	111 120 129 147 207 210 215-216 238
	3		250 330 353 359 366 383-384 414 478
			505 508-509 511 515-524 534-535 537
•			544-545 564 566 571 577 591 598 638
			0

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			663 671 698 714 722 725 727 751 798 851 859 873 876 909 927 949 952 983- 984 1002 1023 1042-1044 1085 1095 1131 1144 1178 1199 1233 1240-1270 1331 1340
fetal liver-spleen	Columbia University	FLS003	64 535 976 1256
fetal liver	Invitrogen	FLV001	8 101 120 138 217 446 468 535 566 580 722 730 749 844 918 943 976 1051 1256 1331
fetal liver	Clontech	FLV004	537 926 1256
fetal muscle	Invitrogen	FMS001	51 111 264 312 369-370 404 417-421 425 535 537 577 598 614 836 857 1141 1208 1268
fetal muscle	Invitrogen	FMS002	537
fetal skin	Invitrogen	FSK001	13-26 32 41 51 89 107 111 147 151 225 264 316 405 422-429 488-494 496 519 534-535 537 566 675 732 859 876- 877 898 947 949-950 963 976 1001 1062 1076 1083 1117 1144 1165 1268 1281
fetal skin	Invitrogen	FSK002	537 812
fetal spleen	BioChain	FSP001	87 549
umbilical cord	BioChain	FUC001	27-33 41 49 151 215 238 248-249 301 316 446 495-503 519 521 534-535 537 582 634 691 877 883 927 944-950 963 976 1001 1075 1142-1143 1171 1218 1243 1308
fetal brain	GIBCO	HFB001	41 49 57 79 87 103 111 120 132-135 138 145 151 188 197 207 215 238 264 271 294 316 367 414 440 446 466 504 513-514 535 542-543 550 564 571 596 635 648-654 675 711-715 722-723 798 832 872 876 883 927 976 1095 1144 1168 1171 1178 1211 1335
macrophage	Invitrogen	HMP001	238
infant brain	Columbia University	IB2002	49-50 77 81 89 105 111 136-138 140 151 161 175-179 185 216-217 264 295 299 308-310 371-373 462 476 504 511- 513 533 537 564 566 571 655-657 662 683 716-720 723 752 790-803 829 832 858-859 876 898 909 949 976 1045- 1047 1076-1087 1090 1093 1116 1122 1144 1209-1213 1225 1233 1256 1319 1341
	Columbia University		41 50 77 104 132 215 238 508 512-513 519 566 655 714 794 918 943 976 1067 1092-1093 1233
infant brain	Columbia University	IBM002	311 472-473 753 1214
infant brain	Columbia University	IBS001	51 111 376 474 790 876 949 1144 1204 1221
lung, fibroblast	Strategene	LFB001	151 316 462 514 534 582 675 939 1131
lung tumor	Invitrogen	LGT002	1-7 41 74 79 94 115 120 138-139 156 215 217 269 280 296 337 374-375 384 404 446 454 475-480 498 514 518-519 522 537 545 564 577 597 653 658 705 721-724 754-756 779 859 868 872-874 876-877 919 927 949 951-952 959 976 1002 1042 1048-1053 1076 1083 1088- 1089 1131 1144-1147 1216-1218 1229

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
		<u> </u>	1293 1311
lymphocytes	ATCC	LPC001	41 74 111 132 151 253 316 446 550
			634 844 927 976 1085 1268
leukocyte	GIBCO	LUC001	8 11 41 74 86 91-98 101 109 111 120
			147 151 212 215 218 238 252 288 312-
			314.316 338 359 408 427 443-447 505
			510 512 514 518 534 545 549-550 561
			564 566 571 577 580 582 587-609 615
			632-638 658-659 698 714 725-728 832
1			836 841 859 866 873-874 882-883 918-
	i e		919 927 943 952 963 976 1042 1076
			1083 1090 1148 1152 1168 1195 1219-
leukocyte	Clontech	LUC003	1220 1224
leukocyte	Cioniech	100003	74 100 215 232 238 339-341 446 545
			657 660 729 873 883 927 952 963 1008
Melanoma from cell	Clontech	MEL004	1042 1116 1120 1149-1150 1215 1222 210 215 238 342 534 545 592 722 873
line ATCC #CRL	Ciontecn	MICLUU4	919 929 939 952 976 1071 1118 1218
1424			1
mammary gland	Invitrogen	MMG001	1235 1245 8-10 40-41 49 73 80 114 138-140 147
maninary granu	TIMITOSCII	IATIAIGOOL	217 250-256 264 297-299 305 377-378
			398 446 481-486 505 512 537 545 549
			571 592 725 730-733 816 829 836 844
			868 873 876-877 898 926 943 951-960
			963 976 995 1034 1042 1048 1054-
			1055 1076 1083 1091 1093 1116-1117
]			1124 1152 1302
induced neuron cells	Strategene	NTD001	39 101 111 138 238 361 1225 1251
	Dau-Sene	1112001	1319
retinoid acid induced	Strategene	NTR001	74 225 976
neuronal cells			
neuronal cells	Strategene	NTU001	129 225 238 304 313 361 657 976
pituitary gland	Clontech	PIT004	976
placenta	Clontech	PLA003	38 976
prostate	Clontech	PRT001	111 188 238 257-258 564 724 961-966
			1067 1095
rectum	Invitrogen	REC001	238 430-431 841 859 868 963 1001
		<u> </u>	1116
salivary gland	Clontech	SAL001	8 151 402 432-433 446 496 868 952
			976 1083 1120 1151 1184
small intestine	Clontech	SIN001	8 101 147 215 259-266 446 462 505
		.,	545 592 660 789 836 866 873 927 952
] .		1,	963 967-978 1042 1120 1152 1223-
		J	1224
skeletal muscle	Clontech	SKM001	238 302 927 943 992 1031
spinal cord	Clontech	SPC001	74 111 132 151 215-216 238 264 267-
			270 343-344 353 379 516 537 566 740
			828 927 976 979-994 1092 1153-1159
		ļ	1225 1250
adult spleen	Clontech	SPLc01	698 859 1042
stomach	Clontech	STO001	210 238 271-272 537 580 705 918 952
<u> </u>			995 1171
thalamus	Clontech	THA002	61 219-220 273-276 312 315 330 596
L			963 996-1007 1059 1093 1160-1162
thymus	Clonetech	THM001	8 120 151 208 221 316-317 353 639
			750 867 874 878-881 927 963 1023
	<u> </u>		1083 1094-1096 1124
thymus	Clontech	THMc02	8 61 114 129 132 210 225 231 306
			317-319 336 340 359 380 398 446 448-
ì		(463 512 519 545 554 587 598 698 724- 725 789 812 836 868 873 927 947 952

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			976 1007 1042 1083 1085 1097-1116
			1122 1147 1177 1226-1229 1234 1311
		<u>.</u> .	1313
thyroid gland	Clontech	THR001	14 41 49 76 94 111 144 151 183 188
			210 217 222 253 264 271 277-286 294
1			320-326 345-352 361 381-382 446 467
İ			483 514 534 549-550 564 578 602 649
			844 882-883 927 950 956 976 1008-
Ì			1028 1076 1083 1117-1120 1142 1163-
ļ		<u> </u>	1175 1230-1238 1308
trachea	Clontech	TRC001	223-225 238 287 353-354 514
}			545 592 611 873 883-884 927
		ļ	952 1029-1031 1042 1151-1152
			1170 1176-1177 1239
uterus	Clontech	UTR001	151 226 288-290 355 537 877
	1		885-886 976 1001 1032-1033
			1232

TABLE 2

SEQ	Accession	Species	Description	Smith-	7%
ID	No.	Í		Waterman	Identity
NO:				Score	
1	B02829	Homo sapiens	Human G protein coupled receptor hRUP5 protein SEQ ID NO:10.	460	100
2	G03564	Homo sapiens	Human secreted protein, SEQ ID NO: 7645.	111	51
3	R26173	Homo sapiens	Part of Major Yo paraneoplastic antigen (CDR62) encoded by clone pY2.	293	76
4	L29536	Homo sapiens	calcium channel L-type alpha 1 subunit	191	65
5	Y94943	Homo sapiens	Human secreted protein clone yt14_1 protein sequence SEQ ID NO:92.	251	50
6	M11507	Homo sapiens	transferrin receptor	120	95
7	AF099100	Homo sapiens	WD-repeat protein 6	1941	93
8	Y92338	Homo sapiens	Human cancer associated antigen precursor from clone NY-REN-45.	245	82
9	G01343	Homo sapiens	Human secreted protein, SEQ ID NO: 5424.	226	91
10	AJ133798	Homo sapiens	copine VII protein	1127	68
11	G02449	Homo sapiens	Human secreted protein, SEQ ID NO: 6530.	584	99
12	X98330	Homo sapiens	ryanodine receptor 2	282	78
13	AL024498	Homo sapiens	dJ417M14.2 (novel serine/threonine-protein kinase (ortholog of mouse and rat MAK (male germ cell-associated kinase))	293	100
14	AF045577	Pan troglodytes	olfactory receptor OR93Ch	191	36
15	G03131	Homo sapiens	Human secreted protein, SEQ ID NO: 7212.	93	39
16	U26595	Rattus norvegicus	prostaglandin F2a receptor regulatory protein precursor	569	89
17	B08918	Homo sapiens	Human secreted protein sequence encoded by gene 28 SEQ ID NO:75.	99	44
18	Y36203	Homo sapiens	Human secreted protein #75.	165	75
19	U15647	Mus musculus	reverse transcriptase	106	40
20	G02701	Homo sapiens	Human secreted protein, SEQ ID NO: 6782.	544	100
21	Y35923	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 172.	1691	100
22	G04030	Homo sapiens	Human secreted protein, SEQ ID NO: 8111.	380	96
23	G02455	Homo sapiens	Human secreted protein, SEQ ID NO: 6536.	123	50
24	AF036329	Homo sapiens	gonadotropin-releasing hormone precursor, second form	284	90
25	G04067	Homo sapiens	Human secreted protein, SEQ ID NO: 8148.	96	32
26	S80119	Rattus sp.	reverse transcriptase homolog	100	34
27	U83303	Homo sapiens	line-1 reverse transcriptase	101	35
28	G03267	Homo sapiens	Human secreted protein, SEQ ID NO: 7348.	135	45

SEQ	Accession	Species	Description	Smith-	1%
ID	No.	apecies	Description	Waterman	1
NO:	1 ****.			Score	Identity
29	G04067	Homo sapiens	Human secreted protein, SEQ ID NO: 8148.	83	42
30	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	116	72
31	G02872 G03371	Homo sapiens	Human secreted protein, SEQ ID NO: 6933. Human secreted protein, SEQ ID NO: 7452.	96	67
32	G03371	Homo sapiens	Human secreted protein, SEQ ID NO: 7452.	58	32
33	Y66688	Homo sapiens	Membrane-bound protein PRO1152.	2457	98
34	Y87071	Homo sapiens	Human secreted protein sequence SEQ ID	348	98
		110mo sapiens	NO:110.	340	1 33
35	U15131	Homo sapiens	p126	182	48
36	Y73464	Homo sapiens	Human secreted protein clone yl4 1 protein	982	90
	1,3,10,1	Tromo suprems	sequence SEQ ID NO:150.	762	,
37	AL133215	Homo sapiens	bA108L7.6 (semaphorin 4G (sema domain,	687	99
			immunoglobulin domain (Ig), transmembrane	1 007	"
			domain (TM) and short cytoplasmic domain))	ŀ	
38	AC067969	amino acids	Homo sapiens ryanodine receptor 1 (skeletal)	386	66
		3338-4088]		
39	AL031588	Homo sapiens	dJ1163J1.1 (mostly supported by GENSCAN,	493	76
		1	FGENES and GENEWISE)	1	1
40	G03628	Homo sapiens	Human secreted protein, SEQ ID NO: 7709.	110	51
41	AF132969	Homo sapiens	CGI-35 protein	228	68
42	Y36268	Homo sapiens	Human secreted protein encoded by gene 45.	220	88
43	X61048	Hydra sp.	mini-collagen	105	35
44	M76546	Helianthus	hydroxyproline-rich protein	110	31
		annuus			
45	U82288	Caenorhabditi	Rac-like GTPase	139	70
		s elegans		1	1
46	G03477	Homo sapiens	Human secreted protein, SEQ ID NO: 7558.	118	58
47	AF090942	Homo sapiens	PRO0657	113	63
48	G03564	Homo sapiens	Human secreted protein, SEQ ID NO: 7645.	90	59
49	AJ005560	Mus	SPR2B protein	72	56
		musculus			<u> </u>
50	G02450	Homo sapiens	Human secreted protein, SEQ ID NO: 6531.	385	98
51	Y91649	Homo sapiens	Human secreted protein sequence encoded by	973	94
	1,000	<u> </u>	gene 60 SEQ ID NO:322.		
52	U93563	Homo sapiens	putative p150	105	38
53	Y55927	Homo sapiens	Human STLK2 protein.	699	85
54	G02607	Homo sapiens	Human secreted protein, SEQ ID NO: 6688.	145	56
55	AB008175	Mus musculus	hepatic nuclear factor 1-beta short form	356	74
54	1460041				<u> </u>
56 57	M68941 AL031600	Homo sapiens	protein-tyrosine phophatase	165	41
58	AE031600 AF011417	Homo sapiens	c390E6.1 (chloride channel 7)	338	76
ەد	Ar01141/	Mus	putative pheromone receptor	143	55
59	AF167320	musculus Mus	ning financia (PD) 12		1
J7	AL 10/320	musculus	zinc finger protein ZFP113	558	68
60	U73036	Homo sapiens	interferon regultory factor 7	262	100
61	X07984	Mus	interferon regultory factor 7 protein-tyrosine kinase	263	96
J1	AU/204	musculus	protein-tyrositie kinase	297	69
62	Y29861	Homo sapiens	Human secreted protein clone cb98 4.	791	98
63	U35376	Homo sapiens	repressor transcriptional factor	485	65
64	AF265555	Homo sapiens	ubiquitin-conjugating BIR-domain enzyme	785	74
7.		110mo suprons	APOLLON	103	′*
65	G03883	Homo sapiens	Human secreted protein, SEQ ID NO: 7964.	88	95
66	AF177390	Manduca	antennal specific membrane protein AMP	274	54
		sexta	mi	1 -17	"
67	AB040800	Homo sapiens	SREB2	614	100
68	AF030027	Equine	24	213	26
· -		herpesvirus 4	- ·	1 213	120
		Homo sapiens	Human secreted protein, SEQ ID NO: 7046.	261	95
69	I G02965	I HOMO Samens		1 201	
69 70	G02965 W75770			1144	98
70	W75770	Homo sapiens	Human oxidoreductase YTFO3.	1144	98 76
	W75770 AB011135	Homo sapiens Homo sapiens	Human oxidoreductase YTFO3. KIAA0563 protein	239	76
70 71	W75770	Homo sapiens	Human oxidoreductase YTFO3.		
70 71	W75770 AB011135	Homo sapiens Homo sapiens Halocynthia	Human oxidoreductase YTFO3. KIAA0563 protein HrPOPK-1	239 813	76 78
70 71 72	W75770 AB011135 AB014885	Homo sapiens Homo sapiens Halocynthia roretzi	Human oxidoreductase YTFO3. KIAA0563 protein	239	76

SEO	Accession	Species	Description	Smith-	1%
ID	No.	Species	Description	Waterman	
NO:	140.		•	Score	Identity
110.	 	musculus		Score	
75	Y00826	Rattus	gp210 (AA 1-1886)	413	84
		norvegicus	SF-10 (1211 1660)	1	"'
76	AF117754	Homo sapiens	thyroid hormone receptor-associated protein	351	54
	}	-	complex component TRAP240		
77	Y38422	Homo sapiens	Human secreted protein.	468	76
78	Y14596	Homo sapiens	Human T-type voltage-gated Ca channel alpha-	1357	99
			1-I (hCavT3).		
79	Y14591	Human	APM-1 protein	767	100
		papillomaviru			•
80	AL137802	s type 68 Homo sapiens	dJ798A10.2 (KIAA0445 protein)	71	- 24
81	AP000383	Arabidopsis	protein arginine N-methyltransferase-like protein	359	65
01	71 000303	thaliana	protein arginine in-meury transferase-like protein	339	63
82	L46815	Mus	DNA binding protein Rc	895	75
		musculus	2 Simoning protein ite	"	/3
83	G01600	Homo sapiens	Human secreted protein, SEQ ID NO: 5681.	315	96
84	Y53886	Homo sapiens	A suppressor of cytokine signalling protein	538	71
			designated HSCOP-6.	1 _	1
85	AB029002	Homo sapiens	KIAA 1079 protein	134	42
86	Y28678	Homo sapiens	Human cw272_7 secreted protein.	325	62
87	Y99368	Homo sapiens	Human PRO1326 (UNQ686) amino acid	156	48
88	AJ225124	Mus	sequence SEQ ID NO:100.	105	105
00	AJ223124	musculus	hyperpolarization-activated cation channel, HAC3	487	95
89	AF177203	Homo sapiens	cerebral cell adhesion molecule	290	56
90	Y28280	Homo sapiens	Human G-protein coupled receptor GRIR-2.	326	79
91	L39891	Homo sapiens	polycystic kidney disease-associated protein	1751	95
92	AF064876	Homo sapiens	ion channel BCNG-1	953	99
93	AF170723	Homo sapiens	protein kinase STK 10	401	53
94	X13292	Trypanosoma	GPI-phospholipase C (AA 1 - 358)	151	37
		brucei	,	ĺ	
95	Y34127	Homo sapiens	Human potassium channel K+Hnov11.	661	99
96	X03638	Rattus	sodium channel protein I (aa 1-2009)	1775	92
07	 	norvegicus			<u> </u>
97 98	AF134213 G00838	Homo sapiens	ubiquitin-specific protease	1995	99
99	AF021935	Homo sapiens Rattus	Human secreted protein, SEQ ID NO: 4919.	213	38
77	AF021933	norvegicus	mytonic dystrophy kinase-related Cdc42-binding kinase	675	48
100	AF279265	Homo sapiens	putative anion transporter 1	867	98
101	AC007878	Homo sapiens	match to nuclear protein, NP220; note: sequence	160	60
			difference at residue 58	100	00
102	U22829	Mus	P2Y purinoceptor	264	42
		musculus			
103	Y45023	Homo sapiens	Human sensory transduction G-protein coupled	516	99
	200 100 0		receptor-B3.		
104	Y94990	Homo sapiens	Human secreted protein vb21_1, SEQ ID NO:20.	787	98
105	Y87342	Homo sapiens	Human signal peptide containing protein HSPP-	343	57
106	AF169312	Homo sapiens	119 SEQ ID NO:119.	212	67
107	AF116657	Homo sapiens	hepatic angiopoietin-related protein PRO1310	212 74	67 52
108	AE000401	Escherichia	sialic acid transporter	74 587	96
200	1111000401	coli	siano aoia nansportor	707	70
109	Y38395	Homo sapiens	Human secreted protein encoded by gene No. 10.	693	100
110	Y78801	Homo sapiens	Hydrophobic domain containing protein clone	182	94
	L _		HP00631 amino acid sequence.		'
111	Z25535	Homo sapiens	nuclear pore complex protein hnup153	464	85
112	Y94939	Homo sapiens	Human secreted protein clone ye90_1 protein	274	51
			sequence SEQ ID NO:84.		
113	AF016365	Homo sapiens	hexokinase 1 isoform td	301	71
114	AC007956	Homo sapiens	unknown	520	75
115	M83738	Homo sapiens	protein-tyrosine phosphatase	251	92
116	AL157952	Homo sapiens	dJ875K15.1.1 (ets homologous factor (ets-	484	91
117	W18084	Homo sapiens	domain transcription factor ESE-3A, isoform 1))	£46	07
	T 44 10004	Tomo sapiens	Human Aurora-2.	546	87

SEQ	Accession	Species	Description	Cmith	1.0/
ID	No.	Species	Description	Smith-	%
NO:	140.			Waterman Score	Identity
118	L41816	Homo sapiens	cam kinase I	407	62
119	AJ006710	Rattus	phosphatidylinositol 3-kinase	627	93
,	12000710	norvegicus	phospitaticy into sitor 3-killase	027	1 33
120	AF026954	Bos taurus	pyruvate dehydrogenase phosphatase regulatory	1646	94
101	020200	 	subunit precursor; PDPr	253	
121	S39392	Homo sapiens	protein tyrosine phosphatase, PTPase {EC 3.1.3.48}	373	68
122	U60805	Homo sapiens	oncostatin-M specific receptor beta subunit	262	88
123	Y44403	Homo sapiens	Human truncated tankyrase-1.	111	35
124	U88167	Caenorhabditi	contains similarity to C2 domains	219	29
	1	s elegans			
125	AF300648	Homo sapiens	guanine nucleotide binding protein beta subunit	693	90
126	AB021861	Mus	apoptosis signal-regulating kinase 2	153	65
120	75021001	musculus	apoptosis signar-regulating Killase 2	133	03
127	AF305210	Homo sapiens	concentrative Na+-nucleoside cotransporter	807	97
			hCNT3	1	
128	M90360	Homo sapiens	protein kinase	220	73
129	D32202	Homo sapiens	alpha 1C adrenergic receptor isoform 2	574	86
130	AF208043	Homo sapiens	IF116b	496	67
131	AF201734	Mus	testis specific serine kinase-3	800	87
132	AF112886	musculus Bos taurus	differentiation enhancing factor 1	159	74
132	AF112886 AJ278314	Homo sapiens	differentiation enhancing factor 1 phospholipase C-beta-1b	554	74 85
134	W74802	Homo sapiens	Human secreted protein encoded by gene 73	1157	87
134	VV /40UZ	110mo sapiens	clone HSOEL25.	1137	0'
135	AB020335	Homo sapiens	Pancreas-specific gene	668	96
136	W80408	Homo sapiens	A secreted protein encoded by clone dt674 2.	866	98
137	AC002563	Homo sapiens	putative RHO/RAC effector protein; 95%	5041	99
138	V06726	1770	similarity to P49205 (PID:g1345860)	001	100
138	Y96736 AB024034	Homo sapiens Arabidopsis	PRO3434, a novel secreted protein.	891	100
137	ABU24U34	thaliana	DNA-damage inducible protein DDI1-like	14/	55
140	W97809	Homo sapiens	Human GTPase regulator GRAF.	248	56
141	Y51557	Homo sapiens	Human PLA2 protein.	125	46
142	AF090113	Rattus	AMPA receptor binding protein	623	93
		norvegicus			
143	W26642	Homo sapiens	Human RECK cancer-inhibiting protein.	641	82
144	U87306	Rattus	transmembrane receptor UNC5H2	578	84
145	AF264014	norvegicus Homo sapiens	scavenger receptor cysteine-rich type 1 protein	727	92
	1 201014	Tromo sapiens	M160 precursor	121	122
146	W63683	Homo sapiens	Human secreted protein 3.	140	40
147	M96264	Homo sapiens	galactose-1-phosphate uridyl transferase	513	81
148	D64014	Escherichia	HrsA	818	90
149	M83316	coli Escherichia	nnnCnn nhosuhohudroken	015	06
147	14103310	coli	pppGpp phosphohydrolase	915	95
150	AL163279	Homo sapiens	homolog to cAMP response element binding and	1261	99
	<u> </u>		beta transducin family proteins		[.]
151	AF179867	Homo sapiens	STE20-like kinase	940	99
152	R95332	Homo sapiens	Tumor necrosis factor receptor 1 death domain	392	61
153	AF151859	Home comies	ligand (clone 3TW).	270	02
154	X66957	Homo sapiens Homo sapiens	CGI-101 protein hexokinase type 1	370 489	92 81
155	Y16355	Homo sapiens	alternatively spliced form	432	92
156	G00857	Homo sapiens	Human sccreted protein, SEQ ID NO: 4938.	349	78
157	AF159455	Mus	zinc finger protein	352	74
		musculus			'
158	L76191	Homo sapiens	interleukin-1 receptor-associated kinase	537	76
159	AP001743	Homo sapiens	putative gene, ankirin like, possible dual	670	98
160	A 1050405	Dotter	specifity Ser/Thr/Tyr kinase domain	556	
160	AJ250425	Rattus norvegicus	Collybistin I	556	74
161	G02885	Homo sapiens	Human secreted protein, SEQ ID NO: 6966.	370	100
			Decided protein, one 10 110, 0700.		

SEQ	Accession	Species	Description	Smith-	T%
JD ID	No.	Species	Description	Waterman	Identity
NO:	140.			Score	lacinity
162	Z22968	Homo sapiens	M130 antigen	610	100
163	AF181121	Homo sapiens	ATP-dopendent Ca2+ pump PMR1	336	92
	AF055636			455	94
164		Homo sapiens	leucine-rich glioma-inactivated protein precursor		
165	AF160798	Rattus norvegicus	calcium transporter CaT1	700	96
166	Y76332	Homo sapiens	Fragment of human secreted protein encoded by gene 38.	327	45
167	Y48607	Homo sapiens	Human breast tumour-associated protein 68.	1072	99
168	AB020741	Mus musculus	NIK-related kinase	197	43
169	AF252293	Homo sapiens	PAR3	596	44
170	U59429	Cricetinae	diacylglycerol kinase eta	481	82
		gen. sp.			
171	AF035268	Homo sapiens	phosphatidylserine-specific phospholipase A1	386	42
172	AF127085	Mus musculus	semaphorin cytoplasmic domain-associated protein 3B	507	82
173	Y27918	Homo sapiens	Human secreted protein encoded by gene No. 123.	653	99
174	G02979	Homo sapiens	Human secreted protein, SEQ ID NO: 7060.	538	97
175	U36488	Mus musculus	embryonic stem cell phosphatase	168	55
176	W95629	Homo sapiens	Homo sapiens secreted protein gene clone	1022	100
		<u> </u>	gm196_4.		<u> </u>
177	AF289023	Homo sapiens	formiminotransferase cyclodeaminase form D	255	93
178	X04936	Homo sapiens	T-cell receptor alpha-chain (413 is 2nd base in codon)	710	99
179	AF127481	Homo sapiens	non-ocogenic Rho GTPase-specific GTP exchange factor	175	80
180	G00978	Homo sapiens	Human secreted protein, SEQ ID NO: 5059.	517	94
181	Y66645	Homo sapiens	Membrane-bound protein PRO1310.	671	96
182	AF110640	Homo sapiens	orphan seven-transmembrane receptor	862	100
183	AB020854	Bos taurus	orphan transporter short splicing variant	766	84
184	AF169691	Homo sapiens	cadherin-like protein VR8	375	38
185	AF126372	Homo sapiens	thyrotropin-releasing hormone degrading ectoenzyme	985	99
106	1 20066	IVana saniana		541	76
186	L20966	Homo sapiens	phosphodiesterase		· · · · · · · · · · · · · · · · · · ·
187	G02920	Homo sapiens	Human secreted protein, SEQ ID NO: 7001.	254	93
188	Y94918	Homo sapiens	Human secreted protein clone dd504_18 protein sequence SEQ ID NO:42.	301	98
189	Y66713	Homo sapiens	Membrane-bound protein PRO1309.	694	100
190	G03244	Homo sapiens	Human secreted protein, SEQ ID NO: 7325.	331	73
191	U36771	Rattus norvegicus	sn-glycerol 3-phosphate acyltransferase	707	92
192	R05935	Homo sapiens	Secreted GPIIb subunit of multiple subunit polypeptide (MSP)GPIIb-IIIa.	157	72
193	M92084	Theileria parva	casein kinase II alpha subunit	364	50
194	Y66645	Homo sapiens	Membrane-bound protein PRO1310.	448	90
195	W95631	Homo sapiens	Homo sapiens secreted protein gene clone	382	49
196	AF255614	Rattus	hj968_2. scaffolding protein SLIPR	680	99
197	AC021640	norvegicus Arabidopsis	putative phosphatidate phosphohydrolase	300	41
		thaliana	· · · · · · · · · · · · · · · · · · ·		
198	AF073967	Mus musculus	olfactory receptor	316	43
	<u> </u>	domesticus	<u> </u>		
199	W01730	Homo sapiens	Human G-protein receptor HPRAJ70.	617	98
200	AF117948	Homo sapiens	pancreas-enriched phospholipase C	625	89
201	AF128625	Homo sapiens	CDC42-binding protein kinase beta	636	94
202	AF117946	Homo sapiens	Link guanine nucleotide exchange factor II	1303	100
203	Y53021	Homo sapiens	Human secreted protein clone qc646_1 protein sequence SEQ ID NO:48.	701	99
			,		1
204	AF227968	Homo sapiens	SH2-B beta signaling protein	182	79

SEQ	Accession	Species	Description	Smith-	1%
D`	No.	1		Waterman	Identity
NO:				Score	1
	~		{ovarian cancer critical region of deletion}		
206	U18315	Sus scrofa	parathyroid receptor .	122	60
207	AF255342	Homo sapiens	putative pheromone receptor V1RL1 long form	170	96
208	S52051	Rattus sp.	neurotransmitter transporter	715	94
209	W63683	Homo sapiens	Human secreted protein 3.	840	99
210	D79992	Homo sapiens	similar to Drosophila photoreceptor cell-specific	541	82
210	D.,,,,,	110mo sapions	protein, calphotin.	341	02
211	AF117948	Homo sapiens	pancreas-enriched phospholipase C	1348	99
212	U81035	Rattus	ankyrin binding cell adhesion molecule	471	69
212	081033	norvegicus	neurofascin	4/1	69
213	AF154846		1	700	
214	AF102777	Homo sapiens Mus	zinc finger protein	798	56
214	AF 102///		FYVE finger-containing phosphoinositide kinase	933	93
216	A7 162202	musculus			<u> </u>
215	AL 163303	Homo sapiens	putative gene containing transmembrane domain	523	89
216	U26595	Rattus	prostaglandin F2a receptor regulatory protein	563	78
	-	norvegicus	precursor	[. <u></u>	
217	G04095	Homo sapiens	Human secreted protein, SEQ ID NO: 8176.	644	98
218	X75756	Homo sapiens	protein kinase C mu	314	81
219	Y66723	Homo sapiens	Membrane-bound protein PRO1100.	770	98
220	D88577	Mus	Kupffer cell receptor	567	40
	<u> </u>	musculus		L	<u>L</u>
221	AF258465	Homo sapiens	OTRPC4	853	100
222	AF021935	Rattus	mytonic dystrophy kinase-related Cdc42-binding	636	96
		norvegicus	kinase	1	
223	AL136527	Homo sapiens	bA215B13.1 (A kinase (PRKA) anchor protein	693	100
		•	11)		"""
224	AB032417	Homo sapiens	WNT receptor Frizzled-4	690	99
225	AF030430	Mus	semaphorin VIa	703	68
		musculus		'**	"
226	AE000218	Escherichia	putative dihydroxyacetone kinase (EC 2.7.1.2)	297	39
		coli	parameter and a only accionic kinacio (EC 2.7.1.2)	***	"
227	AF302150	Homo sapiens	phosphoinositol 3-phosphate-binding protein-2	2080	100
228	AB024573	Mus	GTP-binding like protein 2	265	88
	112021373	musculus	O11 -binding like protein 2	203	00
229	AF122924	Xenopus	Wnt inhibitory factor-1	316	40
		laevis	With initionary factor-1	310	40
230	G03205	Homo sapiens	Human secreted protein, SEQ ID NO: 7286.	229	100
231	X98260	Homo sapiens	M-phase phosphoprotein 11	265	92
232	R92754	Homo sapiens	Human growth differentiation factor-12.		
233	R75111	Homo sapiens	Glycosyl-phosphatidylinositol-specific	682	95
233	K/3111	Homo sapiens	Olycosyl-phosphatidylinositol-specific	290	100
224	37/(0421	-	phospholipase-D.		I
234	W69431	Homo sapiens	Human secreted protein cw1233_3.	235	97
235	Y08686	Homo sapiens	serine palmitoyltransferase, subunit II	859	81
236	AF118275	Homo sapiens	atrophin-related protein ARP	117	37
237	X81466	Mus	Embryo Brain Kinase	460	62
		musculus	, , , , , , , , , , , , , , , , , , , ,		
238	U64857	Caenorhabditi	similar to the BPTI/Kunitz family of inhibitors;	284	33
	-	s elegans	most similar to tissue factor pathway inhibitor		İ
	<u> </u>		precursor (TFPI)]
239	AJ250840	Mus	serine/threonine protein kinase	739	63
		musculus			Į.
240	AJ223472	Mus	transcription elongation factor TFIIS.h	222	38
		musculus	•		1
241	Y94906	Homo sapiens	Human secreted protein clone rb649 3 protein	353	52
			sequence SEQ ID NO:18.		
242	AF169301	Homo sapiens	Na+/sulfate cotransporter SUT-1	591	99
243	L22022	Rattus	orphan transporter v7-3	667	93
		norvegicus		50,	^
244	AF016191	Rattus	potassium channel	1043	98
- + +	711 010191	norvegicus	potassium enaunei	1043	70
245	AF097366	Homo sapiens	cone radium calaium notacium	616	100
246			cone sodium-calcium potassium exchanger	645	98
246 247	Y29868	Homo sapiens	Human secreted protein clone pp325_9.	497	98
	AF180475	Homo sapiens	Not4-Np	188	83
	3717007	1 TT			
248 249	Y17227 AF250910	Homo sapiens Manduca	Human secreted protein (clone ya1-1). death-associated small cytoplasmic leucine-rich	690 182	99

SEQ ID	Accession No.	Species	Description	Smith- Waterman	% Identity
NO:	110.	1	}	Score	Identity
	<u> </u>	sexta	protein SCLP	1	
250	AF192756	Kaposi's	Orf73	134	34
	-	sarcoma-	•	4	}
		associated			
261	AD022604	herpesvirus	MOV	1 200	102
251 252	AB022694 W55045	Homo sapiens Homo sapiens	MOK protein kinase Neural adhesion molecule (ethb0018f2 product).	209 469	83
252 253	L46815	Mus	DNA binding protein Rc	251	100 67
233	L40015	musculus	DIVA bilding protein AC	231	07
254	W68505	Homo sapiens	Human acid sensing ionic channel.	173	82
255	AF070066	Mus	Citron-K kinase	1201	98
	1	musculus		1	
256	G02491	Homo sapiens	Human secreted protein, SEQ ID NO: 6572.	460	100
257	Z12841	Oryctolagus	Phospholipase	368	80
	1	cuniculus		<u> </u>	
258	Y95436	Homo sapiens	Human calcium channel SOC-3/CRAC-2.	1857	99
259	AJ222968	Mus musculus	L-periaxin	430	72
260	AJ250839	Homo sapiens	serine/threonine protein kinase	861	100
261	AJ249977	Homo sapiens	AMP-activated protein kinase gamma 3 subunit	758	98
262	AF141386	Rattus	SLIT-2	198	40
		norvegicus		1.75	
263	AF022859	Homo sapiens	neuropilin-2(a0)	335	62
264	AF160477	Homo sapiens	Ig superfamily receptor LNIR precursor	387	91
265	Y44662	Homo sapiens	Human 14273 G-protein coupled receptor	636	99
	ļ. <u></u>	<u></u>	(GPCR).	<u>]</u>	
266	U27269	Mus	sodium glucose cotransporter	204	56
267	AF124491	musculus	ADE CED	150	-
268	AF124491 AF127389	Homo sapiens Rattus	ARF GTPase-activating protein GIT2 putative taste receptor TR I	159 209	75 39
200	AI 12/309	norvegicus	putative taste receptor TKT	209	39
269	X98296	Homo sapiens	ubiquitin hydrolase	215	95
270	X78482	Streptococcus	Fc-gamma receptor	129	26
	1	pyogenes	,		
271	AB009883	Nicotiana	KED	109	26
		tabacum			
272	AF137367	Mus	VPS10 domain receptor protein SORCS	899	97
273	L34938	musculus Rattus	innotrania alutamata recentor	160	96
213	L34936	norvegicus	ionotropic glutamate receptor	460	86
274	AL022724	Homo sapiens	dJ413H6.1.1 (hamster Androgen-dependent	188	74
	1202721	Tromo supremo	Expressed Protein LIKE PUTATIVE protein)	100	'
	İ	1	(isoform 1)		
275	AF265555	Homo sapiens	ubiquitin-conjugating BIR-domain enzyme	173	94
			APOLLON		
276	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	148	56
277	L40380	Homo sapiens	thyroid receptor interactor	430	61
278	AB046851	Homo sapiens	KIAA1631 protein	283	96
279	AC008075	Arabidopsis thaliana	Contains PF 00069 Eukaryotic protein kinase domain.	157	43
280	M83738	Homo sapiens	protein-tyrosine phosphatase	181	73
281	AK024397	Homo sapiens	unnamed protein product	439	91
282	AF141326	Homo sapiens	RNA helicase HDB/DICE1	497	84
283	AF156530	Mus	ETS-domain transcriptional repressor PE1	605	76
	1	musculus	· · ·	{	
284	Y29336	Homo sapiens	Human secreted protein clone cs756_2 alternate	647	100
	<u> </u>		reading frame protein.		L
285	Y73402	Homo sapiens	Human secreted protein clone yc25_1 protein	300	90
-	1.70		sequence SEQ ID NO:26.		<u> </u>
286	AF016411	Homo sapiens	KCNA3.1B	137	100
287 288	W89253	Homo sapiens	Human ALP.	688	97
288 289	AF112886 AF113131	Bos taurus Homo sapiens	differentiation enhancing factor 1 host cell factor homolog LCP	750 367	96 44
209 290	U52111	Homo sapiens	plexin-related protein	698	100
291	AF026504	Rattus	SPA-1 like protein p1294	603	89

SEQ	Accession	Species	Description	Smith-	1%
ID `	No.	1 -		Waterman	Identity
NO:	1	}		Score	Identity
	1	norvegicus		1 300,0	
292	AF102854	Rattus	membrane-associated guanylate kinase-	124	53
-/-	12 102031	norvegicus	interacting protein 2 Maguin-2	124	33
293	X99211	Drosophila	ubiquitin-specific protease	143	38
275	1777211	melanogaster	disquidit-specific protease	143	30
294	Y94943	Homo sapiens	Human secreted protein clone yt14 1 protein	185	94
-, .	17.5.5	Azomo sapions	sequence SEQ ID NO:92.	103	"4"
295	Y94890	Homo sapiens	Human protein clone HP02798.	108	59
296	AF019767	Homo sapiens	zinc finger protein		
297	Y28568	Homo sapiens		154	96
298	Y94943		Secreted peptide clone bd577_1.	568	84
270	1 94943	Homo sapiens	Human secreted protein clone yt14_1 protein sequence SEQ ID NO:92.	182	97
299	B08906	Homo sapiens	Human secreted protein sequence encoded by	100	1
237	B00700	Hoing sapiens		605	69
300	R58890	Vama assissa	gene 16 SEQ ID NO:63.	1	100
301	AF022859	Homo sapiens	Human-32 cadherin-related molecule.	212	97
302		Homo sapiens	neuropilin-2(a0)	277	100
	Y71124	Homo sapiens	Human mitogenic regulator duox2.	716	97
303	Y44297	Homo sapiens	Human receptor tyrosine kinase.	228	97
304	D32050	Homo sapiens	alanyl-tRNA synthetase	192	80
305	U43586	Homo sapiens	protein kinase related to Raf protein kinases;	428	72
		\	Method: conceptual translation supplied by		
206	D5 4000	 	author		
306	R54872	Homo sapiens	Human H13 viral receptor mutant 4.	280	95
307	D78572	Mus	membrane glycoprotein	199	41
200	1	musculus			
308	AF255614	Rattus	scaffolding protein SLIPR	639	88
		norvegicus		<u> </u>	
309	S79463	Mus sp.	semaphorin homolog=M-Sema F	162	89
310	AF178941	Homo sapiens	ATP-binding cassette sub-family A member 2	736	100
311	U03413	Dictyostelium	calcium binding protein	151	36
		discoideum		1	
312	Y87347	Homo sapiens	Human signal peptide containing protein HSPP-	744	100
	<u> </u>		124 SEQ ID NO:124.		1
313	Z97055	Homo sapiens	dJ388M5.4 (putative GS2 like protein)	789	99
314	AC004010	Homo sapiens	similar to Leucine-rich transmembrane proteins;	197	38
			44% similarity to U42767 (PID:g1736918)		
315	AL021392	Homo sapiens	dJ439F8.2 (supported by GENSCAN and	278	38
	1		GENEWISE)		
316	U70209	Mus	polycystic kidney disease 1 protein	165	38
		musculus			Ĺ <i>'</i>
317	AF109643	Rattus	coxsackie-adenovirus-receptor homolog	223	38
		norvegicus			
318	AF104923	Homo sapiens	putative transcription factor	138	84
319	AF100287	Trypanosoma	activated protein kinase C receptor homolog	141	38
200	-	vivax			<u></u>
320	G00588	Homo sapiens	Human secreted protein, SEQ ID NO: 4669.	125	51
321	Y21591	Homo sapiens	Human secreted protein (clone CC332-33).	459	97
322	D26070	Homo sapiens	human type 1 inositol 1,4,5-trisphosphate	232	97
			receptor	L	
323	Y27918	Homo sapiens	Human secreted protein encoded by gene No.	306	88
	ļ <u></u>		123.	<u> </u>	
324	AF010144	Homo sapiens	neuronal thread protein AD7c-NTP	209	70
325	M19650	Homo sapiens	2',3'-cyclic-nucleotide 3'-phosphodiesterase (EC	214	97
	<u> </u>		3.1.4.37)	l]
326	W80396	Homo sapiens	A secreted protein encoded by clone bp646_10.	140	70
327	X75756	Homo sapiens	protein kinase C mu	540	78
328	G02292	Homo sapiens	Human secreted protein, SEQ ID NO: 6373.	721	99
329	AF168990	Homo sapiens	putative GTP-binding protein	877	99
330	S67984	Homo sapiens	anti-HIV gp120 antibody heavy chain variable	581	80
	1	1	region		1
	X13916	Homo sapiens	LDL-receptor related precursor (AA -19 to 4525)	2823	98
331	1 15.0710				
331 332	Y87330	Homo sapiens	Human signal peptide containing protein HSPP-	l 1127	1 100
	<u> </u>	Homo sapiens	Human signal peptide containing protein HSPP- 107 SEQ ID NO:107.	1127	100
	<u> </u>	Homo sapiens Homo sapiens	Human signal peptide containing protein HSPP- 107 SEQ ID NO:107. HGFH3 Human Growth Factor Homologue 3.	320	98

SEQ	Accession	Species	Description	Smith-	1%
ID	No.			Waterman	Identity
NO:				Score	1
	1		similarity to P49205 (PID:g1345860)	 	
335	Y87347	Homo sapiens	Human signal peptide containing protein HSPP-	1111	67
	j	1	124 SEQ ID NO:124.		}
336	AF006466	Mus	lymphocyte specific formin related protein	193	75
		musculus			1
337	AF265555	Homo sapiens	ubiquitin-conjugating BIR-domain enzyme	632	97
			APOLLON		1
338	Y13443	Homo sapiens	Amino acid sequence of hSlo3-2.	516	100
339	Y07637	Homo sapiens	putative GABA-gated chloride channel	189	100
340	Y05734	Homo sapiens	Human Grb7 effector 2.2412 protein.	2156	99
341	AE000497	Escherichia	L-idonate transcriptional regulator	928	98
	,	coli	2 idonate a ansoriptional rogulator	120	1 20
342	D90855	Escherichia	glycerol-3-phosphate dehydrogenase (EC	769	99
		coli	1.1.99.5) chain A, anaerobic	1 "	"
343	D85613	Escherichia	membrane component	399	100
0 10	203013	coli	incinorate component	"	100
344	M93239	Escherichia	transmembrane protein	232	100
•••	,5225	coli	inaismonistatio protein	232	100
345	M60177	Escherichia	enterobactin	759	99
J 13	14100177	coli	Chicabactin	1 /3/	33
346	D90699	Escherichia	Sensor protein copS (EC 2.7.3).	638	97
2.0	1570077	coli	Solisor protoni copo (EC 2.7.5).	030) "
347	D90843	Escherichia	CapB protein.	552	100
547		coli	Capb protein.	1 332	100
348	M13422	Escherichia	49 kd protein	1193	96
2.0	1113 122	coli	45 kg ploton	1173	100
349	L10328	Escherichia	similar to drug resistance translocases	340	90
547	1 110320	coli	Similar to drug resistance transfocases	340	1 30
350	X69942	Mus	enhancer-trap-locus-1	560	82
330	1000742	musculus	Cinianeer-uap-locus-1	300	02
351	AF239613	Homo sapiens	apamin-sensitive small-conductance Ca2+-	463	80
JJ1	711 257015	Tromo sapiens	activated potassium channel	403	80
352	D90777	Escherichia	3-hydroxybutyryl-CoA dehydrogenase (EC	577	100
JJ2	1550777	coli	1.1.1.157) (b- hydroxybutyryl-CoA	3//	100
		1 00	dehydrogenase) (BhbD).	1	Ï
353	D90863	Escherichia	similar to	311	98
<i>333</i>	250005	coli	Simular to	1 311	100
354	Y52386	Homo sapiens	Human transmembrane protein HP02000.	133	58
355	Y31645	Homo sapiens	Human transport-associated protein-7 (TRANP-	482	55
333	1510.5	Tronto sapiens	7).	402	33
356	Y58637	Homo sapiens	Protein regulating gene expression PRGE-30.	119	51
357	AF119226	Homo sapiens	dual-specificity tyrosine phosphatase YVHI	1788	100
358	Y87219	Homo sapiens	Human secreted protein sequence SEQ ID	165	100
550	10,21,	Troine sapiens	NO:258.	103	100
359	J00132	Homo sapiens	beta-fibrinogen	233	93
360	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	128	70
361	R28916	Homo sapiens	Type III procollagen (prior art).	108	40
362	U16655	Rattus	phospholipase C delta-4	649	65
502	0.0055	norvegicus	риогрионразе С цена-4	U+3	رن [
363	G03119	Homo sapiens	Human secreted protein SEA ID NO. 7200	95	12-
364	U47276	Gallus gallus	Human secreted protein, SEQ ID NO: 7200. chicken brain factor-2	104	42
365	G03789	Homo sapiens			34
366	G04091		Human secreted protein, SEQ ID NO: 7870.	183	65
367		Homo sapiens	Human secreted protein, SEQ ID NO: 8172.	118	46
	X98258	Homo sapiens	M-phase phosphoprotein 9	564	75
368	AL021366	Homo sapiens	cICK0721Q.3 (Kinesin related protein)	3387	99
369	U70932	Peromyscus	reverse transcriptase	92	59
200	1 2000	leucopus		<u> </u>	<u> </u>
370	X86400	Homo sapiens	gamma subunit of sodium potassium ATPase	242	73
	0001=5		like		
371	G03172	Homo sapiens	Human secreted protein, SEQ ID NO: 7253.	165	56
372	U49974	Homo sapiens	mariner transposase	257	55
373	X13916	Homo sapiens	LDL-receptor related precursor (AA -19 to 4525)	21 193	99
374	AF234765	Rattus	serine-arginine-rich splicing regulatory protein	1182	78
	1	norvegicus	SRRP86	1	l
375	U49974	Homo sapiens	mariner transposase	172	55

No. No.	SEQ	Accession	Species	I Description	7 6 7.	T-24
No.		N .	Species	Description	Smith-	%
		140.	· ·			Identity
		C01004	¥7.	**		J
Mass						
Mathematics Mathematics				Human secreted protein, SEQ ID NO: 4750.		
	378	X32574	· ·	GTP binding protein	1456	91
380 309074 Homo sapiens JAK-4p Sign Sig	350					
AB002405 Homo supiens LAK-4p 530 43						
182						37
discoideum dis					530	43
383 G02916 Homo sapiens Human secreted protein, SEQ ID NO; 5975 618 98 98 98 385 AJ245822 Homo sapiens Human secreted protein, SEQ ID NO; 5275 617 93 385 AJ245822 Homo sapiens Type I transmembrane receptor 4560 100 1	382	U64830		protein tyrosine kinase	115	44
385 AJ245822 Homo sapiens Human secreted protein, SEQ ID NO: 5275. 617 93					}	
385 AJ245822 Homo sapiens Open I transmembrane receptor 4560 100			Homo sapiens	Human secreted protein, SEQ ID NO: 6997.	618	98
386 386974 Homo sapiens KIAA0220 12148 98 98 387 G03203 Homo sapiens Human secreted protein, SEQ ID NO: 7284. 142 50 59 59 59 59 59 59 59			Homo sapiens	Human secreted protein, SEQ ID NO: 5275.	617	93
387 G03203 Homo sapiens Human secreted protein, SEQ ID NO: 7284. 142 50 30 388 M040772 Homo sapiens Human secreted protein, SEQ ID NO: 8153. 99 59 59 39 39 39 47293109 Homo sapiens Human secreted protein, SEQ ID NO: 8153. 99 59 59 39 39 47293109 Homo sapiens NFIP2 protein 197 51 197			Homo sapiens	type I transmembrane receptor	4560	100
388 G03203 Homo sapiens Human secreted protein, SEQ ID NO: 7284. 142 50		D86974	Homo sapiens	KIAA0220	2148	98
Human secreted protein, SEQ ID NO: 8153. 99 59 59	387	G03203	Homo sapiens	Human secreted protein, SEO ID NO: 7284.		1
M12140 Homo sapiens envelope protein 197 51	388	G04072				
399 AJ293309 Homo sapiens MFIP protein 461 77 78 79 77 78 79 78 79 78 79 78 79 78 79 78 79 79	389	M12140				
391 Y42751 Homo sapiens Human calcium binding protein 2 (CaBP-2). 181 94	390	AJ293309				
1932 W48351	391					
1939 14442				Human breast cancer related protein RCPD2		I
394 W85607 Homo sapiens Secreted protein clone da228 6. 957 100						
395 395 396 397 398 399 399 391405 399 3	` ` `					
gene 38.						
396 G03930		170332	110illo sapiens		171	34
397	306	G03930	L'ama caniona		1250	
Syndactylus Syndactylus						
398	391	AB032904		dopamine receptor D4	105	35
399 Y91405 Homo sapiens Human secreted protein sequence encoded by gene 2 SEQ ID NO:126. 162 37 37 400 Y29861 Homo sapiens Human secreted protein clone cb98 4. 162 37 37 37 38 39 38 39 38 39 38 39 38 39 38 39 39	300	A 1007700			<u> </u>	<u> </u>
Semilar to rat integral membrane glycoprotein; Semilar to rat perm antigen Semilar to rat sperm antigen Semilar to rat integral membrane glycoprotein Semilar to rat sperm antigen Semilar to rat sperm antigen Semilar to rat sperm antigen Semilar to rat sperm antigen Semilar to rat sperm antigen Semilar to rat sperm antigen Semilar to rat sperm antigen Semilar to rat sperm antigen Semilar to rat sperm antigen Semilar to rat sperm antigen Semilar to rat sperm antigen Semilar to rat sperm antigen Semilar to rat sperm antigen Semilar to ra						
400 Y29861 Homo sapiens Human secreted protein clone cb98 4. 162 37	333	191403	Homo sapiens	Human secreted protein sequence encoded by	1047	92
D87002	400	V20861	Homo conione		1.0	ļ
AF100754						
402	401	D87002	Homo sapiens		527	78
Acoust	402	AE100754	Homo coniona		L	
AF075462 Mus musculus ADP-ribosylation factor-directed GTPasc S45 89						1
Museulus Activating protein isoform b 162 30				ADD -: Land -: Company -: ADD -: Land -: Company -: Land -: Company -: Land -: Company -: Land -: Company -: Land -: L		
Human	707	A1073402			545	89
August A	405	V02897			1.0	
Part Part	705	A92007	L '	pol/env	162	30
406 Y30162 Homo sapiens Human dorsal root receptor 4 hDRR4. 325 72 407 AK022626 Homo sapiens unnamed protein product 2833 99 408 L13802 Homo sapiens ribosmal protein small subunit 264 92 409 Y91600 Homo sapiens Human secreted protein sequence encoded by gene 9 SEQ ID NO:273. 1788 89 410 W88745 Homo sapiens Secreted protein encoded by gene 30 clone HTSEVO9. 2004 99 411 AB043953 Mus musculus Chat-H 2628 82 412 Y86233 Homo sapiens Human secreted protein HNTMX29, SEQ ID NO:148. 1014 92 413 U10542 Pan troglodytes MHC class I A 265 71 414 AF155097 Homo sapiens NY-REN-7 antigen 850 95 415 G03203 Homo sapiens Human secreted protein, SEQ ID NO: 7284 88 48 416 Y57911 Homo sapiens Retinoblastoma binding protein-7sequence 3077 87<					ļ	
AK022626	106	V20162		II		
A08				Human dorsal root receptor 4 hDRR4.		
409 Y91600 Homo sapiens Human secreted protein sequence encoded by gene 9 SEQ ID NO:273. 2004 99						
gene 9 SEQ ID NO:273. 176		1213002	Homo sapiens			1
W88745	409	X 91000	Homo sapiens	Human secreted protein sequence encoded by	1788	89
HTSEV09. HTSEV09. 2628 82	410	17/00/245				
411 AB043953 Mus musculus Chat-H 2628 82 412 Y86233 Homo sapiens Human secreted protein HNTMX29, SEQ ID NO: 1014 1014 92 413 U10542 Pan MHC class I A 265 71 414 AF155097 Homo sapiens NY-REN-7 antigen 850 95 415 G03203 Homo sapiens Human secreted protein, SEQ ID NO: 7284. 88 48 416 Y57911 Homo sapiens Human transmembrane protein HTMPN-35. 266 89 417 W27651 Homo sapiens Secreted protein AT205. 481 60 418 Y76884 Homo sapiens Retinoblastoma binding protein-7sequence. 3077 87 419 AF255559 Notothenia coriiceps alpha tubulin 289 68 420 G01984 Homo sapiens Human secreted protein, SEQ ID NO: 6065. 209 74 421 AL109827 Homo sapiens dJ309K20.2 (acrosomal protein ACR55 (similar to rat sperm antigen 4 (SPAG4))) 112 35	410	W 88 /45	Homo sapiens		2004	99
Musculus	411	17042052			<u> </u>	
412 Y86233 Homo sapiens Human secreted protein HNTMX29, SEQ ID NO:148. 413	411	AB043953	1	Chat-H	2628	82
NO:148. NO:148. 1010 101 1	410	7101000			[Í [
413 U10542 Pan troglodytes MHC class I A 265 71 414 AF155097 Homo sapiens NY-REN-7 antigen 850 95 415 G03203 Homo sapiens Human secreted protein, SEQ ID NO: 7284. 88 48 416 Y57911 Homo sapiens Human transmembrane protein HTMPN-35. 266 89 417 W27651 Homo sapiens Secreted protein AT205. 481 60 418 Y76884 Homo sapiens Retinoblastoma binding protein-7sequence. 3077 87 419 AF255559 Notothenia coriiceps alpha tubulin 289 68 420 G01984 Homo sapiens Human secreted protein, SEQ ID NO: 6065. 209 74 421 AL109827 Homo sapiens dJ309K20.2 (acrosomal protein ACR55 (similar to rat sperm antigen 4 (SPAG4))) 112 35 422 AC008075 Arabidopsis F24J5.4 112 35	412	Y86233	Homo sapiens		1014	92
troglodytes 414 AF155097 Homo sapiens NY-REN-7 antigen 850 95 415 G03203 Homo sapiens Human secreted protein, SEQ ID NO: 7284. 88 48 416 Y57911 Homo sapiens Human transmembrane protein HTMPN-35. 266 89 417 W27651 Homo sapiens Secreted protein AT205. 481 60 418 Y76884 Homo sapiens Retinoblastoma binding protein-7sequence. 3077 87 419 AF255559 Notothenia coriiceps alpha tubulin 289 68 420 G01984 Homo sapiens Human secreted protein, SEQ ID NO: 6065. 209 74 421 AL109827 Homo sapiens dJ309K20.2 (acrosomal protein ACR55 (similar to rat sperm antigen 4 (SPAG4))) 112 35					L	}
414 AF155097 Homo sapiens NY-REN-7 antigen 850 95 415 G03203 Homo sapiens Human secreted protein, SEQ ID NO: 7284. 88 48 416 Y57911 Homo sapiens Human transmembrane protein HTMPN-35. 266 89 417 W27651 Homo sapiens Secreted protein AT205. 481 60 418 Y76884 Homo sapiens Retinoblastoma binding protein-7sequence. 3077 87 419 AF255559 Notothenia coriiceps alpha tubulin 289 68 420 G01984 Homo sapiens Human secreted protein, SEQ ID NO: 6065. 209 74 421 AL109827 Homo sapiens dJ309K20.2 (acrosomal protein ACR55 (similar to rat sperm antigen 4 (SPAG4))) 1146 96 422 AC008075 Arabidopsis F24J5.4 112 35	413	U10542	1 1	MHC class I A	265	71
Homo sapiens Human secreted protein, SEQ ID NO: 7284. 88 48					1]
416 Y57911 Homo sapiens Human transmembrane protein HTMPN-35. 266 89					850	95
416 Y57911 Homo sapiens Human transmembrane protein HTMPN-35. 266 89 417 W27651 Homo sapiens Secreted protein AT205. 481 60 418 Y76884 Homo sapiens Retinoblastoma binding protein-7sequence. 3077 87 419 AF255559 Notothenia coriiceps alpha tubulin 289 68 420 G01984 Homo sapiens Human secreted protein, SEQ ID NO: 6065. 209 74 421 AL109827 Homo sapiens dJ309K20.2 (acrosomal protein ACR55 (similar to rat sperm antigen 4 (SPAG4))) 1446 96 422 AC008075 Arabidopsis F24J5.4 112 35				Human secreted protein, SEQ ID NO: 7284.	88	48
417 W27651 Homo sapiens Secreted protein AT205. 481 60 418 Y76884 Homo sapiens Retinoblastoma binding protein-7sequence. 3077 87 419 AF255559 Notothenia coriiceps alpha tubulin 289 68 420 G01984 Homo sapiens Human secreted protein, SEQ ID NO: 6065. 209 74 421 AL109827 Homo sapiens dJ309K20.2 (acrosomal protein ACR55 (similar to rat sperm antigen 4 (SPAG4))) 1446 96 422 AC008075 Arabidopsis F24J5.4 112 35				Human transmembrane protein HTMPN-35.	266	89
418 Y76884 Homo sapiens Retinoblastoma binding protein-7sequence. 3077 87 419 AF255559 Notothenia coriiceps alpha tubulin 289 68 420 G01984 Homo sapiens Human secreted protein, SEQ ID NO: 6065. 209 74 421 AL109827 Homo sapiens dJ309K20.2 (acrosomal protein ACR55 (similar to rat sperm antigen 4 (SPAG4))) 1446 96 422 AC008075 Arabidopsis F24J5.4 112 35					1.	60
419 AF255559 Notothenia coriiceps alpha tubulin 289 68 420 G01984 Homo sapiens Human secreted protein, SEQ ID NO: 6065. 209 74 421 AL109827 Homo sapiens dJ309K20.2 (acrosomal protein ACR55 (similar to rat sperm antigen 4 (SPAG4))) 1446 96 422 AC008075 Arabidopsis F24J5.4 112 35			Homo sapiens			
Coriiceps 1 20 G01984 Homo sapiens Human secreted protein, SEQ ID NO: 6065. 209 74	419	AF255559				
420 G01984 Homo sapiens Human secreted protein, SEQ ID NO: 6065. 209 74 421 AL109827 Homo sapiens dJ309K20.2 (acrosomal protein ACR55 (similar to rat sperm antigen 4 (SPAG4))) 1446 96 422 AC008075 Arabidopsis F24J5.4 112 35				-		
421 AL 109827 Homo sapiens dJ309K20.2 (acrosomal protein ACR55 (similar to rat sperm antigen 4 (SPAG4))) 1446 96 422 AC008075 Arabidopsis F24J5.4 112 35	420	G01984		Human secreted protein, SEO ID NO: 6065	209	74
to rat sperm antigen 4 (SPAG4))) 422 AC008075 Arabidopsis F24J5.4 112 35						
422 AC008075 Arabidopsis F24J5.4 112 35		1	1		- · · · <u>-</u>	-
	422	AC008075	Arabidopsis		112	35
		1			- 	

SEQ	Accession	Species	Description	Smith-	1%
ID	No.	Species	Besitption	Waterman	Identity
NO:	1	1		Score	Idomity
423	AF231705	Homo sapiens	Alu co-repressor 1	1090	100
424	AF234887	Homo sapiens	FLAMINGO 1	6268	97
425	Y35942	Homo sapiens	Extended human secreted protein sequence, SEQ	1961	99
.25		Tromo supreme	ID NO. 191.	1301	"
426	AB009288	Homo sapiens	N-copine	635	98
427	L12392	Homo sapiens	Huntington's Disease protein	16080	99
428	Y94990	Homo sapiens	Human secreted protein vb21_1, SEQ ID NO:20.	768	98
429	AJ293573	Homo sapiens	zinc finger protein Cezanne	542	87
430	Y84441	Homo sapiens	Amino acid sequence of a human RNA-	2074	100
,,,,	101111	Tromo supieno	associated protein.	2014	100
431	G02850	Homo sapiens	Human secreted protein, SEQ ID NO: 6931.	723	95
432	G04067	Homo sapiens	Human secreted protein, SEQ ID NO: 8148.	73	42
433	AF159296	Lycopersicon	extensin-like protein	613	48
,,,,	1.11.137.270	esculentum	extensin-rike protein	013	1 70
434	W48351	Homo sapiens	Human breast cancer related protein BCRB2.	135	44
435	X73874	Homo sapiens	phosphorylase kinase	3442	97
436	AF161426	Homo sapiens	HSPC308	268	74
437	Y30812	Homo sapiens	Human secreted protein encoded from gene 2.	1055	52
438	G03798	Homo sapiens	Human secreted protein, SEQ ID NO: 7879.	168	56
439	X14766	Homo sapiens	GABA-A receptor alpha 1 subunit	2294	96
440	X02344	Homo sapiens	beta-tubulin	311	95
441	AF168418	Homo sapiens	activating signal cointegrator 1	1882	100
442	L11672	Homo sapiens	zinc finger protein	795	54
443	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	93	26
444	A52140	unidentified	HUMAN NDR	2451	100
445	X98330	Homo sapiens	ryanodine receptor 2	9356	99
446	AF116712	Homo sapiens	PRO2738	227	49
447	AF245447	Homo sapiens	sphingosine kinase type 2 isoform	576	99
448	AF133086	Homo sapiens	membrane-type scrine protease 1	2630	94
449	U87305	Rattus	transmembrane receptor UNC5H1	817	93
747	087303	norvegicus	transmemorane receptor ONC3/11	01/	93
450	AF081249	Homo sapiens	JAW1-related protein MRVIIA long isoform	4568	99
451	AC005498	Homo sapiens	R31665 1	316	62
452	M60235	Homo sapiens	granule membrane protein-140	464	73
453	AB036706	Homo sapiens	intelectin	730	88
454	G00918	Homo sapiens	Human secreted protein, SEQ ID NO: 4999.	263	81
455	Y22634	Homo sapiens	Human cytokine inducible regulatory protein-1	192	67
733	122054	Tionio sapiens	(CIRP-1).	192	07
456	Y36705	Homo sapiens	Fragment of human secreted protein encoded by	106	40
430	130703	Tionio sapiens	gene 62.	100	40
457	N91325	Homo sapiens	DNA encoding human growth hormone receptor.	3282	96
458	M19155	Plasmodium	S-antigen precursor	110	36
.50	"""	falciparum	o anagon productor	110	30
459	Y13377	Homo sapiens	Amino acid sequence of protein PRO257.	509	98
460	Y02693	Homo sapiens	Human secreted protein encoded by gene 44	149	43
.00	102033	Tromo suprens	clone HTDAD22.	149	1 73
461	Y14482	Homo sapiens	Fragment of human secreted protein encoded by	184	54
	1	Long Suprens	gene 17.	107	"
462	Y53005	Homo sapiens	Human secreted protein clone pm749_8 protein	135	47
			sequence SEQ ID NO:16.	-50	["'
463	X84960	Triticum	low molecular weight glutenin	109	33
	.	aestivum		107	ا آ
464	W19919	Homo sapiens	Human Ksr-1 (kinase suppressor of Ras).	1781	85
465	AF189764	Mus	alpha/beta hydrolase-1	502	59
	12 105701	musculus	aipita bota iiy diotaso-1	302	"
466	U93569	Homo sapiens	p40	101	30
467	Y41528	Homo sapiens	Fragment of human secreted protein encoded by	1172	99
	1 11520	1 Tomo Suprems	gene 77.	1172	
468	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	149	52
469	AJ000008	Homo sapiens	PI3-kinase	5832	97
470	X70922	Mus	neurotoxin homologue	118	47
.,,	7.10,22	musculus	noarotoxiii nomotogue	. 10	11
	000505	Homo sapiens	Human secreted protein, SEQ ID NO: 7878.	198	75
471	G03797	HOMB camene			

ID NO: 473 C 474 Y 475 V 476 V 477 Y 478 C 479 A 480 C 481 V 482 G 483 A 484 A 485 G 486 U 487 Y 488 A 489 G	G02313 G07007 W93254 W48351 G02693 G01870 AF102777 G03052 W87701 G03119 AF210651 AF010144 G00637 J15174	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Mus musculus Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	gene 62. Human secreted protein, SEQ ID NO: 6394. Breast cancer associated antigen precursor sequence. Human ESRP1 protein. Human breast cancer related protein BCRB2. Human secreted protein encoded by gene 44 clone HTDAD22. Human secreted protein, SEQ ID NO: 5951. FYVE finger-containing phosphoinositide kinase Human secreted protein, SEQ ID NO: 7133. A human membrane fusion protein designated SYTAX1. Human secreted protein, SEQ ID NO: 7200. NAG18 neuronal thread protein AD7c-NTP Human secreted protein, SEQ ID NO: 4718.	Smith- Waterman Score 328 1013 943 236 202 267 3427 123 221 131 124 343	% Identity 100 97 80 65 60 100 92 53 77 39 59
NO: 473 C 474 Y 475 V 476 V 477 Y 478 C 479 A 480 C 481 V 482 G 483 A 484 A 485 G 486 U 487 Y 488 A 489 G	G02313 707007 W93254 W48351 702693 G01870 AF102777 G03052 W87701 G03119 AF210651 AF010144 G00637 715174	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Mus musculus Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 6394. Breast cancer associated antigen precursor sequence. Human ESRP1 protein. Human breast cancer related protein BCRB2. Human secreted protein encoded by gene 44 clone HTDAD22. Human secreted protein, SEQ ID NO: 5951. FYVE finger-containing phosphoinositide kinase Human secreted protein, SEQ ID NO: 7133. A human membrane fusion protein designated SYTAX1. Human secreted protein, SEQ ID NO: 7200. NAG18 neuronal thread protein AD7c-NTP Human secreted protein, SEQ ID NO: 4718.	328 1013 943 236 202 267 3427 123 221 131 124 343	100 97 80 65 60 100 92 53 77 39 59
473 C 474 Y 475 V 476 V 477 Y 478 C 479 A 480 C 481 V 482 G 483 A 484 A 485 G 487 Y 488 A 489 G	707007 707007	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Mus musculus Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 6394. Breast cancer associated antigen precursor sequence. Human ESRP1 protein. Human breast cancer related protein BCRB2. Human secreted protein encoded by gene 44 clone HTDAD22. Human secreted protein, SEQ ID NO: 5951. FYVE finger-containing phosphoinositide kinase Human secreted protein, SEQ ID NO: 7133. A human membrane fusion protein designated SYTAX1. Human secreted protein, SEQ ID NO: 7200. NAG18 neuronal thread protein AD7c-NTP Human secreted protein, SEQ ID NO: 4718.	328 1013 943 236 202 267 3427 123 221 131 124 343	97 80 65 60 100 92 53 77 39 59
474 Y 475 V 476 V 477 Y 478 C 479 A 480 C 481 V 482 G 483 A 484 A 485 G 486 U 487 Y 488 A 489 G	707007 707007	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Mus musculus Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 6394. Breast cancer associated antigen precursor sequence. Human ESRP1 protein. Human breast cancer related protein BCRB2. Human secreted protein encoded by gene 44 clone HTDAD22. Human secreted protein, SEQ ID NO: 5951. FYVE finger-containing phosphoinositide kinase Human secreted protein, SEQ ID NO: 7133. A human membrane fusion protein designated SYTAX1. Human secreted protein, SEQ ID NO: 7200. NAG18 neuronal thread protein AD7c-NTP Human secreted protein, SEQ ID NO: 4718.	1013 943 236 202 267 3427 123 221 131 124 343	97 80 65 60 100 92 53 77 39 59
474 Y 475 V 476 V 477 Y 478 C 479 A 480 C 481 V 482 G 483 A 484 A 485 G 486 U 487 Y 488 A 489 G	707007 707007	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Mus musculus Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	Breast cancer associated antigen precursor sequence. Human ESRP1 protein. Human breast cancer related protein BCRB2. Human secreted protein encoded by gene 44 clone HTDAD22. Human secreted protein, SEQ ID NO: 5951. FYVE finger-containing phosphoinositide kinase Human secreted protein, SEQ ID NO: 7133. A human membrane fusion protein designated SYTAX1. Human secreted protein, SEQ ID NO: 7200. NAG18 neuronal thread protein AD7c-NTP Human secreted protein, SEQ ID NO: 4718.	1013 943 236 202 267 3427 123 221 131 124 343	97 80 65 60 100 92 53 77 39 59
475 V 476 V 477 Y 478 C 479 A 480 C 481 V 482 G 483 A 484 A 485 G 486 U 487 Y 488 A 489 G	W93254 W48351 W02693 G01870 AF102777 G03052 W87701 G03119 AF210651 AF010144 G00637 U15174	Homo sapiens Homo sapiens Homo sapiens Mus musculus Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	sequence. Human ESRP1 protein. Human breast cancer related protein BCRB2. Human secreted protein encoded by gene 44 clone HTDAD22. Human secreted protein, SEQ ID NO: 5951. FYVE finger-containing phosphoinositide kinase Human secreted protein, SEQ ID NO: 7133. A human membrane fusion protein designated SYTAX1. Human secreted protein, SEQ ID NO: 7200. NAG18 neuronal thread protein AD7c-NTP Human secreted protein, SEQ ID NO: 4718.	943 236 202 267 3427 123 221 131 124 343	80 65 60 100 92 53 77 39 59
476 V 477 Y 478 C 479 A 480 C 481 V 482 G 483 A 484 A 485 G 486 U 487 Y 488 A 489 G	W48351 702693 G01870 WF102777 G03052 W87701 G03119 WF210651 WF010144 G00637 U15174	Homo sapiens Homo sapiens Mus musculus Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	Human ESRP1 protein. Human breast cancer related protein BCRB2. Human secreted protein encoded by gene 44 clone HTDAD22. Human secreted protein, SEQ ID NO: 5951. FYVE finger-containing phosphoinositide kinase Human secreted protein, SEQ ID NO: 7133. A human membrane fusion protein designated SYTAX1. Human secreted protein, SEQ ID NO: 7200. NAG18 neuronal thread protein AD7c-NTP Human secreted protein, SEQ ID NO: 4718.	236 202 267 3427 123 221 131 124 343	65 60 100 92 53 77 39 59
476 V 477 Y 478 C 479 A 480 C 481 V 482 G 483 A 484 A 485 G 486 U 487 Y 488 A 489 G	W48351 702693 G01870 WF102777 G03052 W87701 G03119 WF210651 WF010144 G00637 U15174	Homo sapiens Homo sapiens Mus musculus Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	Human breast cancer related protein BCRB2. Human secreted protein encoded by gene 44 clone HTDAD22. Human secreted protein, SEQ ID NO: 5951. FYVE finger-containing phosphoinositide kinase Human secreted protein, SEQ ID NO: 7133. A human membrane fusion protein designated SYTAX1. Human secreted protein, SEQ ID NO: 7200. NAG18 neuronal thread protein AD7c-NTP Human secreted protein, SEQ ID NO: 4718.	236 202 267 3427 123 221 131 124 343	65 60 100 92 53 77 39 59
477 Y 478 C 479 A 480 C 481 V 482 G 483 A 484 A 485 G 487 Y 488 A 489 G	702693 601870 AF102777 603052 V87701 603119 AF210651 AF010144 600637 715174	Homo sapiens Mus musculus Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	Human secreted protein encoded by gene 44 clone HTDAD22. Human secreted protein, SEQ ID NO: 5951. FYVE finger-containing phosphoinositide kinase Human secreted protein, SEQ ID NO: 7133. A human membrane fusion protein designated SYTAX1. Human secreted protein, SEQ ID NO: 7200. NAG18 neuronal thread protein AD7c-NTP Human secreted protein, SEQ ID NO: 4718.	202 267 3427 123 221 131 124 343	53 77 39 59
478 C 479 A 480 C 481 V 482 G 483 A 484 A 485 G 486 U 487 Y 488 A 489 G	G01870 AF102777 G03052 V87701 G03119 AF210651 AF010144 G00637 J15174	Homo sapiens Mus musculus Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	clone HTDAD22. Human secreted protein, SEQ ID NO: 5951. FYVE finger-containing phosphoinositide kinase Human secreted protein, SEQ ID NO: 7133. A human membrane fusion protein designated SYTAX1. Human secreted protein, SEQ ID NO: 7200. NAG18 neuronal thread protein AD7c-NTP Human secreted protein, SEQ ID NO: 4718.	267 3427 123 221 131 124 343	100 92 53 77 39 59
480 C 481 V 482 G 483 A 484 A 485 G 486 U 487 Y 488 A 489 G	AF102777 603052 V87701 603119 AF210651 AF010144 600637 V15174	Mus musculus Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 5951. FYVE finger-containing phosphoinositide kinase Human secreted protein, SEQ ID NO: 7133. A human membrane fusion protein designated SYTAX1. Human secreted protein, SEQ ID NO: 7200. NAG18 neuronal thread protein AD7c-NTP Human secreted protein, SEQ ID NO: 4718.	3427 123 221 131 124 343	92 53 77 39 59
480 C 481 V 482 G 483 A 484 A 485 G 486 U 487 Y 488 A 489 G	AF102777 603052 V87701 603119 AF210651 AF010144 600637 V15174	Mus musculus Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	FYVE finger-containing phosphoinositide kinase Human secreted protein, SEQ ID NO: 7133. A human membrane fusion protein designated SYTAX1. Human secreted protein, SEQ ID NO: 7200. NAG18 neuronal thread protein AD7c-NTP Human secreted protein, SEQ ID NO: 4718.	3427 123 221 131 124 343	92 53 77 39 59
480 C 481. V 482 G 483 A 484 A 485 G 486 U 487 Y 488 A 489 G	603052 V87701 603119 G7210651 G7010144 G00637 J15174	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 7133. A human membrane fusion protein designated SYTAX1. Human secreted protein, SEQ ID NO: 7200. NAG18 neuronal thread protein AD7c-NTP Human secreted protein, SEQ ID NO: 4718.	123 221 131 124 343	53 77 39 59
481. V 482 G 483 A 484 A 485 G 486 U 487 Y 488 A 489 G	V87701 503119 JF210651 JF010144 500637 J15174 776167	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	A human membrane fusion protein designated SYTAX1. Human secreted protein, SEQ ID NO: 7200. NAG18 neuronal thread protein AD7c-NTP Human secreted protein, SEQ ID NO: 4718.	221 131 124 343	77 39 59
481. V 482 G 483 A 484 A 485 G 486 U 487 Y 488 A 489 G	V87701 503119 JF210651 JF010144 500637 J15174 776167	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	A human membrane fusion protein designated SYTAX1. Human secreted protein, SEQ ID NO: 7200. NAG18 neuronal thread protein AD7c-NTP Human secreted protein, SEQ ID NO: 4718.	221 131 124 343	77 39 59
482 G 483 A 484 A 485 G 486 U 487 Y 488 A 489 G	G03119 G10651 G10144 G10637 G15174 G16167	Homo sapiens Homo sapiens Homo sapiens Homo sapiens	SYTAX1. Human secreted protein, SEQ ID NO: 7200. NAG18 neuronal thread protein AD7c-NTP Human secreted protein, SEQ ID NO: 4718.	131 124 343	39
483 A 484 A 485 G 486 U 487 Y 488 A 489 G	AF210651 AF010144 600637 U15174	Homo sapiens Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 7200. NAG18 neuronal thread protein AD7c-NTP Human secreted protein, SEQ ID NO: 4718.	124 343	59
483 A 484 A 485 G 486 U 487 Y 488 A 489 G	AF210651 AF010144 600637 U15174	Homo sapiens Homo sapiens Homo sapiens	NAG18 neuronal thread protein AD7c-NTP Human secreted protein, SEQ ID NO: 4718.	124 343	59
484 A 485 G 486 U 487 Y 488 A 489 G	AF010144 500637 515174 576167	Homo sapiens Homo sapiens	neuronal thread protein AD7c-NTP Human secreted protein, SEQ ID NO: 4718.	343	
485 G 486 U 487 Y 488 A 489 G	00637 115174 776167	Homo sapiens	Human secreted protein, SEQ ID NO: 4718.		טכו
486 U 487 Y 488 A 489 G	76167		and societed proton, 3DQ ID NO. 4/18.	129	70
487 Y 488 A 489 G	76167	-20mo supiens	BCL2/adenovirus E1B 19kD-interacting protein	149	
488 A 489 G		,	3	149	73
488 A 489 G		Homo sapiens	Human secreted protein encoded by gene 44.	627	1 100
489 G	J275213	Homo sapiens	stabilin-1	1244	91
	03798	Homo sapiens	Human secreted protein, SEQ ID NO: 7879.	313	65
490 L	12392	Homo sapiens	Huntington's Disease protein	16081	
	03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	197	100
	03799	Homo sapiens	laminin-binding protein	228	66
	115174	Homo sapiens	BCL2/adenovirus E1B 19kD-interacting protein	128	41
,,,,		riomo supiciis	3	128	41
494 Y	02693	Homo sapiens	Human secreted protein encoded by gene 44	197	67
1	02033	Tiomo suprems	clone HTDAD22.	197	07
495 A	C005175	Homo sapiens	R31449 3	889	94
	03786	Homo sapiens	Human secreted protein, SEQ ID NO: 7867.	229	61
	B030237	Canis	D4 dopamine receptor	90	48
		familiaris	БУ ворание гесерго	90	40
498 G	02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	228	65
	70935	Peromyscus	reverse transcriptase	213	52
		maniculatus	10.030 datistriptabe	213	1
500 U	48508	Homo sapiens	skeletal muscle ryanodine receptor	26406	99
501 G	03371	Homo sapiens	Human secreted protein, SEQ ID NO: 7452.	105	58
	F119851	Homo sapiens	PRO1722	156	62
	F113685	Homo sapiens	PRO0974	116	50
	79458	Homo sapiens	WW domain binding protein-2	322	59
	/29651	Homo sapiens	Human secreted protein CD124_3.	608	55
	/85459	Homo sapiens	Secreted protein encoded by clone dh1135 9.	986	70
	86265	Homo sapiens	Human secreted protein HUSXE77, SEQ ID	115	33
1			NO:180.	113	33
508 A	L160175	Homo sapiens	bA243J16.3 (similar to MYLK (myosin, light	184	92
1			polypeptide kinase))	104	92
509 U	43360	Peromyscus	reverse transcriptase	97	62
		maniculatus		<i>,</i> ,	02
510 G	03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	117	63
	79092	Homo sapiens	Human secreted protein dn740_3.	1058	100
	F010144	Homo sapiens	neuronal thread protein AD7c-NTP	205	64
		Homo sapiens	GRIP1 protein	2151	100
		Drosophila	CG6393 gene product	259	42
		melanogaster	Berra bronger		74 .
515 Z1	17206	Xenopus	p46XIEg22	128	40
-		laevis	h	120	+∪
516 AI		Homo sapiens	large tumor suppressor 1	1766	94
		Homo sapiens	Human secreted protein, SEQ ID NO: 7878.	92	40
		Homo sapiens	HSPC249	444	98
		Homo sapiens	cytochrome c-like polypeptide	318	50
		Plasmodium	pval	170	61
		vivax	h.m.	1/0	01

No. No. Homo sapiens Human secreted protein, SEQ ID NO: 7871. 159 59	SEQ	Accession	Species	Description	Smith-	1%
No.			operies	Description	1	J
521 API21857 Homo sapiens Human secreted protein, SEQ ID NO: 7871. 159 39 522 AFI21857 Homo sapiens Homo sapiens 1259 40 524 W88627 Homo sapiens Human secreted protein encoded by gene 94 clone 253 73 524 W88627 Homo sapiens Homo sapiens Human secreted protein encoded by gene 94 clone 253 73 525 AFI19851 Homo sapiens Human secreted protein, SEQ ID NO: 6788. 70 45 526 V27761 Homo sapiens Human secreted protein, SEQ ID NO: 6788. 70 45 528 U47924 Homo sapiens Human secreted protein, SEQ ID NO: 7884. 111 80 521 G03203 Homo sapiens Human secreted protein, SEQ ID NO: 7844. 84 45 531 G04067 Homo sapiens Human secreted protein, SEQ ID NO: 7848. 192 65 533 G03268 Homo sapiens Human secreted protein, SEQ ID NO: 7844. 182 48 534 AP166286 Homo sapiens		110.				lucinity
API21857 Home sapiens Sorting pexish 7 Society Society Home sapiens Society Home sapiens Society Home sapiens Society Home sapiens Society Home sapiens Society Home sapiens Society Home sapiens Society Home sapiens Society Home sapiens Society Home sapiens Society Home sapiens Society Home sapiens Society Home sapiens Society Home sapiens Society Home sapiens Human secreted protein encoded by gene 94 clone 253 73 Society Soc		G03790	Homo saniens	Human secreted protein SEO ID NO: 7971		50
						_1
Secreted protein encoded by gene 94 clone						
Section				Segreted protein angulad by some 04 along		
1926 192776				HPMBQ32.	253	13
Section Sect				PRO1722	162	57
Section Sect				Human secreted protein encoded by gene No. 47.	154	57
Section Sect			Homo sapiens	Human secreted protein, SEQ ID NO: 6788.	70	45
G03203				I	1112	86
Signature				Human secreted protein, SEQ ID NO: 8144.	84	45
G03267 Homo sapiens Human secreted protein, SEQ ID NO: 7348. 75 29					111	60
San					92	
AF068286 Homo sapiens HDCMD38P 861 100				Human secreted protein, SEQ ID NO: 7348.	75	29
1935 1007707					182	48
1936 G01955 Homo sapiens Human secreted protein, SEQ ID NO: 6036. 484 75 537 AF219322 Gallus gailus Gin-induced kinase 206 53 53 538 AF2135022 Homo sapiens Industrial Medical Residue Industrial Residue			Homo sapiens		861	100
Sample		U07707	Homo sapiens	epidermal growth factor receptor substrate	228	60
AF135022				Human secreted protein, SEQ ID NO: 6036.	484	75
193 G03267 Homo sapiens Human secreted protein, SEQ ID NO: 7348. 141 59			Gallus gallus	qin-induced kinase	206	53
AF016430 Caenorinabditi s elegans Caenorinabditi s elegans Sail Sail Sail AC003093 Homo sapiens OXYSTEROL-BINDING PROTEIN; 45% 408 66 66 66 66 67 67 67 6					128	100
540 AF016430 Caenorhabditical selegans contains similarity to a BR-C/TTK domain selegans 39 541 AC003093 Homo sapiens OXYSTEROL-BINDING PROTEIN; 45% similarity to P22059 (PID:g129308) 408 66 542 M29487 Homo sapiens integrin alpha subunit precursor 517 81 543 AF102530 Mus musculus olfactory receptor F3 327 73 544 Y73431 Homo sapiens human secreted protein clone yb136_1 protein 386 100 545 AE004833 Pseudomonas aeruginosa probable TonB-dependent receptor 279 42 546 G03793 Homo sapiens Human secreted protein, SEQ ID NO: 7874. 264 53 547 Y69192 Homo sapiens Human secreted protein, SEQ ID NO: 7874. 264 53 548 Y91493 Homo sapiens Human secreted protein sequence encoded by gene 43 SEQ ID NO: 166. 100 176 100 549 G01571 Homo sapiens Human secreted protein, SEQ ID NO: 5652. 777 99 550 AF044588 Homo sapiens	539		Homo sapiens	Human secreted protein, SEQ ID NO: 7348.	141	59
541 AC003093 Homo sapiens OXYSTEROL-BINDING PROTEIN; 45% similarity to P22059 (PID:g129308) 408 66 542 M29487 Homo sapiens integrin alpha subunit precursor 517 81 543 AF102530 Mus musculus olfactory receptor F3 327 73 544 Y73431 Homo sapiens Human secreted protein clone yb186_1 protein sequence SEQ ID NO:84. 386 100 545 AE004833 Pseudomonas aeruginosa Human secreted protein, SEQ ID NO:7874. 264 53 546 G03793 Homo sapiens Probable TonB-dependent receptor 279 42 547 Y69192 Homo sapiens A human monocyte-macrophage apolipoprotein B receptor protein. 1772 67 548 Y91493 Homo sapiens Protein regulating cytokinesis regulating cytokinesis regulating cytokinesis regulating special protein sequence encoded by gene 43 SEQ ID NO: 4562. 176 100 549 G01571 Homo sapiens Protein regulating cytokinesis regulating cytokinesis regulating cytokinesis regulating cytokinesis regulating regulating cytokinesis regulating cytokinesis regulating regulating regulating regulating regulating regulating regulating regulating regulating regulating regulation regulating regulating regulation regulating re	540	AF016430			853	39
542 M29487 Homo sapiens integrin alpha subunit precursor 517 81 543 AF102530 Mus musculus offactory receptor F3 327 73 544 Y73431 Homo sapiens Human secreted protein clone yb186_1 protein serulence. SEQ ID No.34. 386 100 545 AE004833 Pseudomonas aeruginosa Probable TonB-dependent receptor 279 42 546 G03793 Homo sapiens Human secreted protein, SEQ ID NO: 7874. 264 53 547 Y69192 Homo sapiens Human secreted protein, SEQ ID NO: 7874. 264 53 548 Y91493 Homo sapiens Human secreted protein sequence encoded by gene 43 SEQ ID NO: 166. 176 100 549 G01571 Homo sapiens Protein regulating cytokinesis 1; PRC1 1953 88 551 Y29332 Homo sapiens Fluman secreted protein, SEQ ID NO: 5652. 777 99 552 X98330 Homo sapiens Human secreted protein, SEQ ID NO: 5652. 777 99 553 Y42782 Homo sapiens	541	AC003093			408	66
543 AF102530 Mus musculus olfactory receptor F3 327 73 544 Y73431 Homo sapiens Human secreted protein clone yb186_1 protein sequence SEQ ID NO:84. 100 545 AE004833 Pseudomonas acruginosa 279 42 546 G03793 Homo sapiens Human secreted protein, SEQ ID NO: 7874. 264 53 547 Y69192 Homo sapiens A human monocyte-macrophage apolipoprotein 1772 67 548 Y91493 Homo sapiens Human secreted protein sequence encoded by gene 43 SEQ ID NO:166. 176 100 549 G01571 Homo sapiens Fluman secreted protein, SEQ ID NO: 5652. 777 99 550 AP044588 Homo sapiens Protein regulating cytokinesis 1; PRC1 1953 88 551 Y29332 Homo sapiens Protein regulating cytokinesis 1; PRC1 1953 88 552 X98330 Homo sapiens Human UE Band #331 protein. 684 95 553 Y42782 Homo sapiens RING-H2 1468 100	542	M29487	Homo saniens		517	81
544 Y73431 Homo sapiens sequence SEQ ID NO:84. 386 100 545 AE004833 Pseudomonas sequence SEQ ID NO:84. 279 42 546 G03793 Homo sapiens Human secreted protein, SEQ ID NO: 7874. 264 53 547 Y69192 Homo sapiens Human secreted protein, SEQ ID NO: 7874. 264 53 548 Y91493 Homo sapiens Human secreted protein sequence encoded by gene 43 SEQ ID NO: 166. 176 100 549 G01571 Homo sapiens Human secreted protein, SEQ ID NO: 5652. 777 99 550 AF044588 Homo sapiens Protein regulating cytokinesis 1; PRC1 1953 88 551 Y29332 Homo sapiens Protein regulating cytokinesis 1; PRC1 1953 88 552 X98330 Homo sapiens Tyanodine receptor 2 24621 99 553 Y42782 Homo sapiens Human UC Band #331 protein. 684 95 554 AB025258 Mus granuphilin-a 501 41 555 <td>543</td> <td>AF102530</td> <td>Mus</td> <td></td> <td></td> <td></td>	543	AF102530	Mus			
AE004833	544	Y73431			386	100
Section Sect	545	AE004833			279	42
Second Process	546	G03793		Human secreted protein CEO ID NO. 7974	264	63
B receptor protein. 1772 07					L =	
Separate SEQ ID NO:166 Separate Sepa			<u> </u>	B receptor protein.		
Section Sect	_		1	gene 43 SEQ ID NO:166.		100
Sequence						1
Sequence						1
State	351	Y29332	-		1224	94
State		X98330	Homo sapiens		24621	99
AB025258	553	Y42782	Homo sapiens	Human UC Band #331 protein.	684	95
AJ010346 Homo sapiens RING-H2 1468 100	554	AB025258	Mus		501	
S56 W92388 Homo sapiens Human TR-interacting protein S239a. 538 92	555	A 1010346		PING-H2	1460	100
Section Sect						
AF117756						
559 G02872 Homo sapiens Human secreted protein, SEQ ID NO: 6953. 319 68 560 D86214 Mus musculus Ca2÷ dependent activator protein for secretion musculus 1010 93 561 AF187325 Canis familiaris melanoma antigen 287 55 562 AJ001981 Homo sapiens OXA1L 2512 99 563 Z17238 Rattus glutamate receptor subtype delta-1 338 66 564 W30638 Homo sapiens Partial human 7-transmembrane receptor HAPO167 protein. 371 100 565 AC005620 Homo sapiens R33590_i 467 97 566 Y99358 Homo sapiens Human PRO1772 (UNQ834) amino acid sequence SEQ ID NO:63. 1138 78 567 AL031177 Homo sapiens dJ889M15.3 (novel protein) 1002 58	558			thyroid hormone receptor-associated protein		
D86214 Mus Ca2÷ dependent activator protein for secretion 1010 93	550	G02872	Home resises		210	-
musculus						
Familiaris Fam					1010	93
562 AJ001981 Homo sapiens OXA1L 2512 99 563 Z17238 Rattus norvegicus glutamate receptor subtype delta-1 338 66 564 W30638 Homo sapiens Partial human 7-transmembrane receptor HAPO167 protein. 371 100 565 AC005620 Homo sapiens R33590_1 467 97 566 Y99358 Homo sapiens Human PRO1772 (UNQ834) amino acid sequence SEQ ID NO:63. 1138 78 567 AL031177 Homo sapiens dJ889M15.3 (novel protein) 1002 58	561	AF187325	1	melanoma antigen	287	55
Sequence SEQ ID NO:63	562	AJ001981		OXAII.	2512	90
norvegicus	563		Rattus			
HAPO167 protein. 100	564	W30638		Partial human 7-transmembrane recentor		100
566 Y99358 Homo sapiens Human PRO1772 (UNQ834) amino acid 1138 78 567 AL031177 Homo sapiens dJ889M15.3 (novel protein) 1002 58				HAPO167 protein.		
567 AL031177 Homo sapiens dJ889M15.3 (novel protein) 1002 58			<u> </u>	sequence SEQ ID NO:63.		78
568 AF151043 Homo sapiens HSPC209 798 100						
	568	AF151043	Homo sapiens	HSPC209	798	100

WO 01/57188

SEQ	Accession	Species	Description	Smith-	1%
ID	No.	operies	Beschpiton	Waterman	Identity
NO:	1	1		Score	Identity
569	AF097518	Homo sapiens	liver-specific transporter	231	100
570	AB035698	Homo sapiens	Misshapen/NIK-related kinase MINK-1	1532	100
571	Y07096	Homo sapiens	Colon cancer associated antigen precursor	1064	100
371	107030	Tionio sapiens	sequence.	1064	100
572	AL031177	Homo sapiens		1 505	1
573			dJ889M15.3 (novel protein)	735	55
	Y66639	Homo sapiens	Membrane-bound protein PRO290.	254	45
574	AB037108	Homo sapiens	seven transmembrane domain orphan receptor	1883	99
575	D43949	Homo sapiens	This gene is novel.	836	100
576	Y48596	Homo sapiens	Human breast tumour-associated protein 57.	108	50
577	G00352	Homo sapiens	Human secreted protein, SEQ ID NO: 4433.	141	75
578	R95913	Homo sapiens	Neural thread protein.	140	65
579	AK025116	Homo sapiens	unnamed protein product	201	70
580	Y86473	Homo sapiens	Human gene 52-encoded protein fragment, SEQ ID NO:388.	77	70
581	AF196779	Homo sapiens	JM10 protein	450	100
582	AF188706	Homo sapiens	g20 protein	330	98
583	AB030234	Canis	D4 dopamine receptor	64	56
		familiaris		} ~~	33
584	G02621	Homo sapiens	Human secreted protein, SEQ ID NO: 6702.	345	90
585	AL096828	Homo sapiens	dJ963E22.1 (Novel protein similar to NY-REN-2	268	85
			Antigen)	200	1 05
586	Y30819	Homo sapiens	Human secreted protein encoded from gene 9.	235	35
587	G00357	Homo sapiens	Human secreted protein, SEQ ID NO: 4438.	132	56
588	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	182	79
589	AF235017	Mus	2P1 protein	764	80
505	10 255017	musculus	21 1 protein	704	00
590	W88627	Homo sapiens	Secreted protein encoded by gene 94 clone	329	81
	1100027	Tionio sapiens	HPMBQ32.	329	01
591	Y30709	Homo sapiens	Amino acid sequence of a human secreted	110	43
371	130,05	110mo sapiens	protein.	110	43
592	Y53875	Homo sapiens	A human seven transmembrane signal transducer	1369	92
-,-	1 2000	Ixomo supiens	polypeptide.	1309	32
593	Y53051	Homo sapiens	Human secreted protein clone dd119_4 protein	1112	97
-,-	1	Tromo dapieno	sequence SEQ ID NO:108.	1112	31
594	Y27658	Homo sapiens	Human secreted protein encoded by gene No. 92.	763	79
595	G03798	Homo sapiens	Human secreted protein, SEQ ID NO: 7879.	156	58
596	AF151110	Mus	COP1 protein	2215	95
	T	musculus	COLI protein	2213	93
597	G03786	Homo sapiens	Human secreted protein, SEQ ID NO: 7867.	157	65
598	AF192499	Mus	putative secreted protein, 3EQ 1D NO. 7807.	143	40
370	11112499	musculus	putative secreted protein 231037	143	40
599	AF119855	Homo sapiens	PRO1847	226	
600·	G02872	Homo sapiens		236	76
601	Y00295	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	212	73
	1.510.1551	 	Human secreted protein encoded by gene 38.	567	88
602	AF184971	Homo sapiens	class II cytokine receptor ZCYTOR7	2015	74
603 604	AF061936	Homo sapiens	diacylglycerol kinase iota	773	96
004	AL096828	Homo sapiens	dJ963E22.1 (Novel protein similar to NY-REN-2	1333	93
606	AD022104	177	Antigen)		
605	AB033106	Homo sapiens	KIAA1280 protein	3915	100
606	X75756	Homo sapiens	protein kinase C mu	3916	99
607	D86983	Homo sapiens	similar to D.melanogaster peroxidasin(U11052)	5758	99
608	W69341	Homo sapiens	Secreted protein of clone CG279_1.	1377	99
609	W88627	Homo sapiens	Secreted protein encoded by gene 94 clone	339	82
	V		HPMBQ32.		
610	Y27868	Homo sapiens	Human secreted protein encoded by gene No.	116	62
611	AE202626	Tioma	107.	21.64	100
611	AF202636	Homo sapiens	angiopoietin-like protein PP1158	2164	100
612	AF090944	Homo sapiens	PRO0663	218	82
613	Y02693	Homo sapiens	Human secreted protein encoded by gene 44	195	59
614	1107053	Datter	clone HTDAD22.	150	
614	M87053	Rattus	lens membrane protein	450	84
615	AC004222	norvegicus	EDA 42 15	162	
	AC004232	Homo sapiens	FPM315	163	37
616	G01984	Homo sapiens	Human secreted protein, SEQ ID NO: 6065.	205	79

SEQ	Accession	Species	Description	Smith-	%
ID NO:	No.			Waterman Score	Identity
617	Y91524	Homo sapiens	Human secreted protein sequence encoded by gene 74 SEQ ID NO:197.	821	99
618	AJ245621	Homo sapiens	CTL2 protein	2258	99
619	Y76198	Homo sapiens	Human secreted protein encoded by gene 75.	108	64
620	AF067864	Homo sapiens	transferrin receptor 2 alpha	3922	94
621	D90721	Escherichia coli	Transmembrane protein dppC	573	90
622	W75858	Homo sapiens	Human secretory protein of clone CS752-3.	730	100
623	Y94982	Homo sapiens	Human secreted protein vb12_1, SEQ ID NO:4.	733	100
624	AF034745	Mus musculus	LNXp80	637	83
625	U42580	Paramecium	Pro-rich, IPPPNMSLPLS (3x)	94	46
		bursaria Chlorella			
		virus 1		-	
626	U79260	Homo sapiens	unknown	194	70
627	R95913	Homo sapiens	Neural thread protein.	99	50
628	G03450	Homo sapiens	Human secreted protein, SEQ ID NO: 7531.	427	100
629	Y36281	Homo sapiens	Human secreted protein encoded by gene 58.	590	100
630	Y02693	Homo sapiens	Human secreted protein encoded by gene 44	165	76
631	G02139		clone HTDAD22.		
632	U16996	Homo sapiens	Human secreted protein, SEQ ID NO: 6220.	268	96
633	AF121857	Homo sapiens	protein tyrosine posphatase	351	80
634	AF121857 AF283772	Homo sapiens	sorting nexin 7	2019	100
		Homo sapiens	similar to Homo sapiens ribosomal protein L10 encoded by GenBank Accession Number L25899	340	77
635	Y07090	Homo sapiens	Renal cancer associated antigen precursor sequence.	277	64
636	AB013382	Homo sapiens	DUSP6	414	76
637	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	315	71
638	M95762	Rattus norvegicus	GABA transporter	924	89
639	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	219	60
640	Y01400	Homo sapiens	Secreted protein encoded by gene 18 clone HNHFO29.	137	79
641	AC008075	Arabidopsis thaliana	F24J5.4	121	33
642	W74824	Homo sapiens	Human secreted protein encoded by gene 96 clone HAQBK61.	615	62
643	AB015982	Homo sapiens	serine/threonine kinase	485	98
644	Y25806	Homo sapiens	Human secreted protein fragment encoded from gene 23.	162	46
645	AF122904	Homo sapiens	membrane protein DAP10	474	100
646	AF233323	Homo sapiens	Fas-associated phosphatase-1	200	38
647	W48804	Homo sapiens	Homo sapiens clone BK158_1 protein.	1203	99
648	AF257330	Homo sapiens	COBW-like protein	1440	98
649	Y36203	Homo sapiens	Human secreted protein #75.	233	73
650	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	173	78
651	Y32199	Homo sapiens	Human receptor molecule (REC) encoded by Incyte clone 2022379.	1012	100
652	AB032909	Hylobates agilis	dopamine receptor D4	122	32
653	AK021848	Homo sapiens	unnamed protein product	186	69
654	W73411	Homo sapiens	Human secreted protein encoded by Gene No. 15.	57	37
655	L22455	Rattus norvegicus	mu opioid receptor	116	34
656	G03112	Homo sapiens	Human secreted protein, SEQ ID NO: 7193.	110	45
657	G02345	Homo sapiens	Human secreted protein, SEQ ID NO: 6426.	459	97
658	W88627	Homo sapiens	Secreted protein encoded by gene 94 clone HPMBQ32.	291	75
659	G02832	Homo sapiens	Human secreted protein, SEQ ID NO: 6913.	134	65
660	Y91423	Homo sapiens	Human secreted protein sequence encoded by	333	96
		Ll	gene 11 SEQ ID NO:144.	l	

SEQ	Accession	Species	Description	Smith-	%
ID	No.	1		Waterman	Identity
NO:		 		Score	<u> </u>
661	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	168	68
662	Y53886	Homo sapiens	A suppressor of cytokine signalling protein	375	43
	Wasaai	+,	designated HSCOP-6.		1.00
663	W75771	Homo sapiens	Human GTP binding protein APD08.	629	100
664	AL096770	Homo sapiens	bA150A6.2 (novel 7 transmembrane receptor	480	55
			(rhodopsin family) (olfactory receptor like)	ł	
665	AB037734	Homo sapiens	protein (hs6M1-21))	978	100
666	W82841		KIAA1313 protein		96
667	W82841	Homo sapiens	Human cerebral protein-1.	192	84
668	AB030184	Homo sapiens	Human cerebral protein-1.	182	87
000	AB030184	Mus musculus	contains transmembrane (TM) region and ATP	757	68
669	AB032919	Hylobates	binding region dopamine receptor D4	05	127
009	AB032919	muelleri	dopamine receptor D4	85	37
670	AF107295	Rattus	outer membrane protein	746	81
070	AF 107293	norvegicus	outer memorane protein	/40	81
671	Z33642	Homo sapiens	leukocyte surface protein	394	93
672	W85608	Homo sapiens	Secreted protein clone du410 5.		
				261	91
673 674	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	106	48
	AL035587	Homo sapiens	dJ475N16.4 (KJAA0240)	2388	99
675	Y59668	Homo sapiens	Secreted protein 108-005-5-0-C1-FL.	1134	53
676	G03797	Homo sapiens	Human secreted protein, SEQ ID NO: 7878.	174	74
677	AF026954	Bos taurus	pyruvate dehydrogenase phosphatase regulatory subunit precursor; PDPr	1013	95
678	L11625	Mus	receptor protein-tyrosine kinase	545	96
		musculus	·		
679	AL031427	Homo sapiens	dJ167A19.3 (novel protein)	745	100
680	AJ133430	Mus	olfactory receptor	528	77
	<u>i</u>	musculus		1	
681	G02532	Homo sapiens	Human secreted protein, SEQ ID NO: 6613.	179	70
682	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	336	76
683	Y94943	Homo sapiens	Human secreted protein clone yt14_1 protein sequence SEQ ID NO:92.	118	100
684	U43360	Peromyscus	reverse transcriptase	100	37
		maniculatus		1	1
685	G00885	Homo sapiens	Human secreted protein, SEQ ID NO: 4966.	162	60
686	AK001518	Homo sapiens	unnamed protein product	590	100
687	G01982	Homo sapiens	Human secreted protein, SEQ ID NO: 6063.	718	100
688	Y92241	Homo sapiens	Human cancer associated antigen precursor (MO-REN-46).	2405	99
		'			1
689	AC024792	Caenorhabditi s elegans	contains similarity to TR:P78316	423	36
		s elegans	contains similarity to TR:P78316		
689 690	AC024792 Y27868	1	contains similarity to TR:P78316 Human secreted protein encoded by gene No. 107.	423 183	36
		s elegans	contains similarity to TR:P78316 Human secreted protein encoded by gene No. 107.		
690 691	Y27868 Y56514	s elegans Homo sapiens	contains similarity to TR:P78316 Human secreted protein encoded by gene No. 107. Human Jurkat cell clone P2-15 AIM10 longest ORF protein sequence.	183	81
690 691 692	Y27868 Y56514	s elegans Homo sapiens	contains similarity to TR:P78316 Human secreted protein encoded by gene No. 107. Human Jurkat cell clone P2-15 AIM10 longest ORF protein sequence.	183	81
690 691 692	Y27868 Y56514	s elegans Homo sapiens Homo sapiens	contains similarity to TR:P78316 Human secreted protein encoded by gene No. 107. Human Jurkat cell clone P2-15 AIM10 longest ORF protein sequence. Human secreted protein encoded by gene No. 79.	183	81
690	Y27868 Y56514	s elegans Homo sapiens Homo sapiens Homo sapiens Homo sapiens	contains similarity to TR:P78316 Human secreted protein encoded by gene No. 107. Human Jurkat cell clone P2-15 AIM10 longest ORF protein sequence.	183 180 1539	88
690 691 692 693	Y27868 Y56514 Y27795 Y36268	s elegans Homo sapiens Homo sapiens Homo sapiens	contains similarity to TR:P78316 Human secreted protein encoded by gene No. 107. Human Jurkat cell clone P2-15 AIM10 longest ORF protein sequence. Human secreted protein encoded by gene No. 79. Human secreted protein encoded by gene 45. ribosomal protein L35	183 180 1539 428 308	81 88 99 98 89
690 691 692 693 694	Y27868 Y56514 Y27795 Y36268 U12465 Y45272	s elegans Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	contains similarity to TR:P78316 Human secreted protein encoded by gene No. 107. Human Jurkat cell clone P2-15 AIM10 longest ORF protein sequence. Human secreted protein encoded by gene No. 79. Human secreted protein encoded by gene 45. ribosomal protein L35 Human secreted protein encoded from gene 16.	183 180 1539 428 308 1517	81 88 99 98 89 99
690 691 692 693 694 695	Y27868 Y56514 Y27795 Y36268 U12465	s elegans Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	contains similarity to TR:P78316 Human secreted protein encoded by gene No. 107. Human Jurkat cell clone P2-15 AIM10 longest ORF protein sequence. Human secreted protein encoded by gene No. 79. Human secreted protein encoded by gene 45. ribosomal protein L35 Human secreted protein encoded from gene 16. TANK binding kinase TBK1 Human secreted protein encoded by gene 44	183 180 1539 428 308	81 88 99 98 89
690 691 692 693 694 695 696	Y27868 Y56514 Y27795 Y36268 U12465 Y45272 AF191838	s elegans Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	contains similarity to TR:P78316 Human secreted protein encoded by gene No. 107. Human Jurkat cell clone P2-15 AIM10 longest ORF protein sequence. Human secreted protein encoded by gene No. 79. Human secreted protein encoded by gene 45. ribosomal protein L35 Human secreted protein encoded from gene 16. TANK binding kinase TBK1 Human secreted protein encoded by gene 44 clone HTDAD22. Human signal peptide containing protein HSPP-	183 180 1539 428 308 1517 1242	81 88 99 98 89 99 98
690 691 692 693 694 695 696 697	Y27868 Y56514 Y27795 Y36268 U12465 Y45272 AF191838 Y02693	s elegans Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	contains similarity to TR:P78316 Human secreted protein encoded by gene No. 107. Human Jurkat cell clone P2-15 AIM10 longest ORF protein sequence. Human secreted protein encoded by gene No. 79. Human secreted protein encoded by gene 45. ribosomal protein L35 Human secreted protein encoded from gene 16. TANK binding kinase TBK1 Human secreted protein encoded by gene 44 clone HTDAD22. Human signal peptide containing protein HSPP- 57 SEQ ID NO:57. Human SCAD family molecule HSFM-1, SEQ	183 180 1539 428 308 1517 1242 275	81 88 99 98 89 99 98 75
690 691 692 693 694 695 696 697 698	Y27868 Y56514 Y27795 Y36268 U12465 Y45272 AF191838 Y02693 Y87280 Y97999	s elegans Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	contains similarity to TR:P78316 Human secreted protein encoded by gene No. 107. Human Jurkat cell clone P2-15 AIM10 longest ORF protein sequence. Human secreted protein encoded by gene No. 79. Human secreted protein encoded by gene 45. ribosomal protein L35 Human secreted protein encoded from gene 16. TANK binding kinase TBK1 Human secreted protein encoded by gene 44 clone HTDAD22. Human signal peptide containing protein HSPP57 SEQ ID NO:57. Human SCAD family molecule HSFM-1, SEQ ID NO:1.	183 180 1539 428 308 1517 1242 275 576	81 88 99 98 89 99 98 75 90
690 691 692 693 694 695 696 697 698 699	Y27868 Y56514 Y27795 Y36268 U12465 Y45272 AF191838 Y02693 Y87280 Y97999	s elegans Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	contains similarity to TR:P78316 Human secreted protein encoded by gene No. 107. Human Jurkat cell clone P2-15 AIM10 longest ORF protein sequence. Human secreted protein encoded by gene No. 79. Human secreted protein encoded by gene 45. ribosomal protein L35 Human secreted protein encoded from gene 16. TANK binding kinase TBK1 Human secreted protein encoded by gene 44 clone HTDAD22. Human signal peptide containing protein HSPP57 SEQ ID NO:57. Human SCAD family molecule HSFM-1, SEQ ID NO:1. putative serine/threonine protein kinase	183 180 1539 428 308 1517 1242 275 576 729	88 99 98 89 99 98 75 90 99
690 691 692 693 694 695 696 697 698 699 700 701	Y27868 Y56514 Y27795 Y36268 U12465 Y45272 AF191838 Y02693 Y87280 Y97999 AJ006701 AF209198	s elegans Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	contains similarity to TR:P78316 Human secreted protein encoded by gene No. 107. Human Jurkat cell clone P2-15 AIM10 longest ORF protein sequence. Human secreted protein encoded by gene No. 79. Human secreted protein encoded by gene 45. ribosomal protein L35 Human secreted protein encoded from gene 16. TANK binding kinase TBK1 Human secreted protein encoded by gene 44 clone HTDAD22. Human signal peptide containing protein HSPP-57 SEQ ID NO:57. Human SCAD family molecule HSFM-1, SEQ ID NO:1. putative serine/threonine protein kinase zinc finger protein 277	183 180 1539 428 308 1517 1242 275 576 729 610 2357	88 99 98 89 99 98 75 90 99 79 100
690 691 692 693 694 695 696 697 698 699	Y27868 Y56514 Y27795 Y36268 U12465 Y45272 AF191838 Y02693 Y87280 Y97999	s elegans Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	contains similarity to TR:P78316 Human secreted protein encoded by gene No. 107. Human Jurkat cell clone P2-15 AIM10 longest ORF protein sequence. Human secreted protein encoded by gene No. 79. Human secreted protein encoded by gene 45. ribosomal protein L35 Human secreted protein encoded from gene 16. TANK binding kinase TBK1 Human secreted protein encoded by gene 44 clone HTDAD22. Human signal peptide containing protein HSPP57 SEQ ID NO:57. Human SCAD family molecule HSFM-1, SEQ ID NO:1. putative serine/threonine protein kinase	183 180 1539 428 308 1517 1242 275 576 729	88 99 98 89 99 98 75 90 99
690 691 692 693 694 695 696 697 698 699 700 701 702	Y27868 Y56514 Y27795 Y36268 U12465 Y45272 AF191838 Y02693 Y87280 Y97999 AJ006701 AF209198 AJ298841	s elegans Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Mus musculus	contains similarity to TR:P78316 Human secreted protein encoded by gene No. 107. Human Jurkat cell clone P2-15 AIM10 longest ORF protein sequence. Human secreted protein encoded by gene No. 79. Human secreted protein encoded by gene 45. ribosomal protein L35 Human secreted protein encoded from gene 16. TANK binding kinase TBK1 Human secreted protein encoded by gene 44 clone HTDAD22. Human signal peptide containing protein HSPP57 SEQ ID NO:57. Human SCAD family molecule HSFM-1, SEQ ID NO:1. putative serine/threonine protein kinase zinc finger protein 277 torsinA protein	183 180 1539 428 308 1517 1242 275 576 729 610 2357 709 -	88 99 98 89 99 98 75 90 99 79 100 45
690 691 692 693 694 695 696 697 698 699 700 701 702 703	Y27868 Y56514 Y27795 Y36268 U12465 Y45272 AF191838 Y02693 Y87280 Y97999 AJ006701 AF209198 AJ298841	s elegans Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	contains similarity to TR:P78316 Human secreted protein encoded by gene No. 107. Human Jurkat cell clone P2-15 AIM10 longest ORF protein sequence. Human secreted protein encoded by gene No. 79. Human secreted protein encoded by gene 45. ribosomal protein L35 Human secreted protein encoded from gene 16. TANK binding kinase TBK1 Human secreted protein encoded by gene 44 clone HTDAD22. Human signal peptide containing protein HSPP57 SEQ ID NO:57. Human SCAD family molecule HSFM-1, SEQ ID NO:1. putative serine/threonine protein kinase zinc finger protein 277 torsinA protein	183 180 1539 428 308 1517 1242 275 576 729 610 2357 709 -	88 99 98 89 99 98 75 90 99 79 100 45
690 691 692 693 694 695 696 697 698 699 700 701 702	Y27868 Y56514 Y27795 Y36268 U12465 Y45272 AF191838 Y02693 Y87280 Y97999 AJ006701 AF209198 AJ298841	s elegans Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Mus musculus	contains similarity to TR:P78316 Human secreted protein encoded by gene No. 107. Human Jurkat cell clone P2-15 AIM10 longest ORF protein sequence. Human secreted protein encoded by gene No. 79. Human secreted protein encoded by gene 45. ribosomal protein L35 Human secreted protein encoded from gene 16. TANK binding kinase TBK1 Human secreted protein encoded by gene 44 clone HTDAD22. Human signal peptide containing protein HSPP57 SEQ ID NO:57. Human SCAD family molecule HSFM-1, SEQ ID NO:1. putative serine/threonine protein kinase zinc finger protein 277 torsinA protein	183 180 1539 428 308 1517 1242 275 576 729 610 2357 709 -	88 99 98 89 99 98 75 90 99 79 100 45

SEQ	Accession	Species	Description	Smith-	%
ID	No.			Waterman	Identity
NO:				Score	
706	G02501	Homo sapiens	Human secreted protein, SEQ ID NO: 6582.	125	58
707	R95326	Homo sapiens	Tumor necrosis factor receptor 1 death domain	121	95
			ligand (clone 2DD).		<u> </u>
708	G03002	Homo sapiens	Human secreted protein, SEQ ID NO: 7083.	125	39
709	Y96202	Homo sapiens	IkappaB kinase (IKK) binding protein, Y2H56.	516	98
710	M63577	Saccharomyc	SFP1	131	59
	17000001	es cerevisiae			L
711	AB026291	Rattus	acetoacetyl-CoA synthetase	467	85
712	D21211	norvegicus	(PTP D 4 C)	1260	
713	AF044033	Homo sapiens Marmota	protein tyrosine phosphatase (PTP-BAS, type 3)	368	44
/13	Ar044033	marmota	olfactory receptor	615	83
714	G03561	Homo sapiens	Human secreted protein, SEQ ID NO: 7642.	251	100
715	AB033062	Homo sapiens	KIAA1236 protein	1380	100
716	G00577	Homo sapiens	Human secreted protein, SEQ ID NO: 4658.	80	73
717	Y96864	Homo sapiens	SEQ. ID. 37 from WO0034474.	835	99
718	AJ243396	Homo sapiens	voltage-gated sodium channel beta-3 subunit	234	100
719	U47334	Homo sapiens	similar to chicken gamma aminobutyric acid	578	99
			receptor beta4 subunit	""	"
720	AB020598	Homo sapiens	peptide transporter 3	1096	100
721	Y53886	Homo sapiens	A suppressor of cytokine signalling protein	570	74
	ł	1.	designated HSCOP-6.		'
722	J05046	Homo sapiens	insulin receptor-related receptor	6787	100
723	AF001958	Ambystoma	electrogenic Na+ bicarbonate cotransporter;	111	41
		tigrinum	NBC		
724	AF127084	Mus	semaphorin cytoplasmic domain-associated	5253	94
		musculus	protein 3A	ł	ł
725	X54673	Homo sapiens	GABA transporter	3114	99
726	AF016191	Rattus	potassium channel	370	100
60.5	1000000	norvegicus			
727	AB029559	Rattus	BATI	139	35
728	Y28503	norvegicus	THORNE II		
729	AJ011415	Homo sapiens	HGFH3 Human Growth Factor Homologue 3.	2186	97
730	Z93096	Homo sapiens Homo sapiens	plexin-B1/SEP receptor bK390B3.1 (manic fringe (Drosophila)	729	56
730	293090	Homo sapiens	homolog)	142	68
731	Z10062	Homo sapiens	cDNA encoding a human vanilloid receptor	675	99
,	210002	Tionio sapiens	homologue Vanilrep1.	0/3) 33
732	AF161382	Homo sapiens	HSPC264	492	94
733	AB029033	Homo sapiens	KIAA1110 protein	3826	99
734	AE000493	Escherichia	putative transport protein	592	97
		coli	Parameter Program	""	1 "
735	AL033379	Homo sapiens	dJ417O22.2 (novel 7 transmembrane receptor	2173	99
	1		(rhodopsin family) protein similar to high-		
			affinity lysophosphatidic acid receptor homolog)		
736	AF132599	Homo sapiens	RANTES factor of late activated T lymphocytes-	245	56
	<u> </u>		1	<u> </u>	
737	X55019	Homo sapiens	acetylcholine receptor delta subunit	883	99.
738	X91906	Homo sapiens	voltage-gated chloride ion channel	1978	100
739	AB026116	Homo sapiens	organic anion transporter 4	1444	98
740	D00570	Mus	open reading frame (196 AA)	83	24
7/1	W/02626	musculus	II	110	10
741 742	W03626	Homo sapiens	Human thyrotropin GPR N-terminal sequence.	118	40
743	U66059	Homo sapiens	V_segment translation product	614	100
744	AF119815 X16663	Homo sapiens	G-protein-coupled receptor	2751	99
744	W67838	Homo sapiens	haematopoietic lineage cell protein (AA 1-486)	148	93
143	1 440/030	Homo sapiens	Human secreted protein encoded by gene 32	448	95
746	W57260	Homo sapiens	clone HLTCJ63.	2414	100
747	W21578	Homo sapiens	Human semaphorin Y. Alzheimer's disease protein encoded by DNA	2414 . 968	100
, , ,	11213/6	Tomo sapiens	from plasmid pGCS2232.	200	65
748	Y94935	Homo sapiens	Human secreted protein clone yd218 1 protein	622	100
		110mo sapiens	sequence SEQ ID NO:76.	<i>322</i>	100
749	AL022238	Homo sapiens	dJ1042K10.5 (novel protein)	314	85
777					

Score Score Score Score Score Score Score	SEQ	Accession	Species	Description	Smith-	1%
	ID NO:	No.				Identity
1753 Y48586 Homo sapiens Human breast tumour-associated protein 47. 2527 99 754 A)272207 Homo sapiens Human breast tumour-associated protein 47. 2527 97 755 M85183 Rattus Native Grotein-coupled receptor 92 649 100 756 AF190501 Homo sapiens Human secreted protein encoded by gene 43 461 87 757 Y02692 Homo sapiens Human secreted protein encoded by gene 43 461 87 758 Z22535 Homo sapiens Human secreted protein encoded by gene 43 461 87 759 R04932 Homo sapiens Human secreted protein encoded by gene 175 1217 99 750 W74902 Homo sapiens Human secreted protein encoded by gene 175 1217 99 751 G03706 Homo sapiens Human secreted protein encoded by gene 175 1217 99 751 ARC026976 Homo sapiens Human secreted protein encoded by gene 175 1217 99 752 Homo sapiens Human secreted protein encoded by gene 175 1217 99 753 ARC026992 Homo sapiens Human secreted protein encoded by gene 175 1217 99 754 ARC026992 Homo sapiens Human secreted protein encoded by gene 175 1217 99 755 ARC026992 Homo sapiens Human secreted protein encoded by gene 175 1217 99 756 AF230378 Homo sapiens Human secreted protein encoded by gene 175 1217 99 757 AF230378 Homo sapiens Human secreted protein encoded by gene 175 1217 99 758 AF230378 Homo sapiens Human secreted protein encoded by gene 175 1217 99 758 AF230378 Homo sapiens Human secreted protein encoded by gene 175 1217 99 759 AF06815 Homo sapiens Human secreted protein encoded by gene 175 1217 99 750 Y09945 Rattus Human secreted protein encoded by gene 175 1217 99 751 AF226731 Homo sapiens Human secreted protein encoded by gene 10 1169 100 757 AF125791 Homo sapiens Human secreted protein encoded by gene 10 1169 100 750 Y09945 Rattus Human secreted protein encoded by gene 10 109 100 750 Y09	751	AB025258	1 '	granuphilin-a	773	41
	752	Y52386	Homo sapiens	Human transmembrane protein HP02000.	900	99
	753	Y48586	Homo sapiens		2527	99
	754	AJ272207		putative G protein-coupled receptor 92	694	100
	755	M85183	Rattus		979	68
			norvegicus		ļ	1
Colone HTADX17. Colone HTA	756	AF190501	Homo sapiens		388	71
	757	Y02692	Homo sapiens		461	87
39 responsible for binding the target.	758	Z22535	Homo sapiens	ALK-3	439	98
Human secreted protein encoded by gene 175 1217 99	759	R04932	Homo sapiens		564	97
AB020676 Homo sapiens KIAA0869 protein 4433 99 AK026992 Homo sapiens Unnamed protein product 2285 39 AK026992 Homo sapiens Glucocorticoid receptor AF-1 coactivator-1 573 100 AF28066 Mus	760	W74902	Homo sapiens	Human secreted protein encoded by gene 175	1217	99
AB020676 Homo sapiens KIAA0869 protein 4433 99	761		Homo sapiens	Human secreted protein, SEQ ID NO: 7787.	223	88
	762	AB020676			4433	99
	763	AK026992		unnamed protein product	2285	99
AF268066 Mus musculus netrin 4 musculus musculus musculus musculus interleukin-1 delta 309 45 45 45 45 45 45 45 4	764	AF173358			573	100
AF230378 Mus musculus Mus musculus Mus musculus Mus musculus Mus musculus Mus Musculus Mus Musculus Mus Musculus Muscu	765	AF268066		netrin 4	2019	89
Mus	766	Y48585	Homo sapiens	Human breast tumour-associated protein 46.	1169	89
Macro Macr	767	AF230378		interleukin-1 delta	309	45
Y09945 Ratus norvegicus Putative integral membrane transport protein A58 S0	768	AF121975		odorant receptor S18	268	62
Norwegicus	769	AB008515	Homo sapiens	RanBPM	611	57
	770	Y09945		putative integral membrane transport protein	458	50
NOV/plexin-Al protein 1821 98	771	AF226731	Homo sapiens	AD026	688	99
AB025258	772	Y27132	Homo sapiens		1384	100
AB025258	773	X87832	Homo sapiens	NOV/plexin-A1 protein	1821	98
	774	AB025258	1	granuphilin-a	500	41
1977 G02493 Homo sapiens Human secreted protein, SEQ ID NO: 6574. 191 68	775	AF125101	Homo sapiens		232	93
R03301	776	G02815	Homo sapiens	Human secreted protein, SEQ ID NO: 6896.	314	95
AL357374 Homo sapiens bA353C18.2 (novel protein) 232 100	777	G02493	Homo sapiens	Human secreted protein, SEQ ID NO: 6574.	191	68
AF100346	778		Homo sapiens	Sequence of pre-human atrial natriuretic peptide.	213	45
3 subunit 3 subunit 103 52	779	AL357374			232	100
Protein Prot	780	AF100346	Homo sapiens		1434	89
AF084464 Rattus norvegicus REM2 Rattus Rattus norvegicus REM2 Rattus Remosapiens REM2 Rattus Remosapiens REM2 Rattus Remosapiens REM2 Rattus Remosapiens Remos	781	Y19566	Homo sapiens		103	52
	782		Homo sapiens		1098	
W49042	783	AF084464		GTP-binding protein REM2	141	30
86 Y91870 Homo sapiens Human apoptosis related protein. 547 100 87 Y71062 Homo sapiens Human membrane transport protein, MTRP-7. 1062 94 88 AF117754 Homo sapiens thyroid hormone receptor-associated protein complex component TRAP240 8684 98 89 AL049569 Homo sapiens dJ37C10.3 (novel ATPase) 2848 96 90 AF151848 Homo sapiens CGI-90 protein 745 96 91 Y08639 Homo sapiens nuclear orphan receptor ROR-beta 1421 95 92 Y41706 Homo sapiens Human PRO381 protein sequence. 644 99 93 AF121228 Homo sapiens thyroid hormone receptor-associated protein complex component TRAP95 1037 100 94 G04072 Homo sapiens Human secreted protein, SEQ ID NO: 8153. 124 62 95 Y69384 Homo sapiens Amino acid sequence of a 14274 receptor 119 100	784	W49042			2693	99
86 Y91870 Homo sapiens Human apoptosis related protein. 547 100 87 Y71062 Homo sapiens Human membrane transport protein, MTRP-7. 1062 94 88 AF117754 Homo sapiens thyroid hormone receptor-associated protein complex component TRAP240 8684 98 89 AL049569 Homo sapiens dJ37C10.3 (novel ATPase) 2848 96 90 AF151848 Homo sapiens CGI-90 protein 745 96 91 Y08639 Homo sapiens nuclear orphan receptor ROR-beta 1421 95 92 Y41706 Homo sapiens Human PRO381 protein sequence. 644 99 93 AF121228 Homo sapiens thyroid hormone receptor-associated protein complex component TRAP95 1037 100 94 G04072 Homo sapiens Human secreted protein, SEQ ID NO: 8153. 124 62 95 Y69384 Homo sapiens Amino acid sequence of a 14274 receptor 119 100	785	AF238381	Homo sapiens	PTOV1	1904	91
87 Y71062 Homo sapiens Human membrane transport protein, MTRP-7. 1062 94 88 AF117754 Homo sapiens thyroid hormone receptor-associated protein complex component TRAP240 8684 98 89 AL049569 Homo sapiens dJ37C10.3 (novel ATPase) 2848 96 90 AF151848 Homo sapiens CGI-90 protein 745 96 91 Y08639 Homo sapiens nuclear orphan receptor ROR-beta 1421 95 92 Y41706 Homo sapiens Human PRO381 protein sequence. 644 99 93 AF121228 Homo sapiens thyroid hormone receptor-associated protein complex component TRAP95 1037 100 94 G04072 Homo sapiens Human secreted protein, SEQ ID NO: 8153. 124 62 95 Y69384 Homo sapiens Amino acid sequence of a 14274 receptor 119 100	786	Y91870			547	100
Complex component TRAP240 Complex component TRAP240	787	Y71062		Human membrane transport protein, MTRP-7.	1062	94
	788	AF117754	Homo sapiens		8684	98
	789	AL049569	Homo sapiens		2848	96
192Y41706Homo sapiensHuman PRO381 protein sequence.64499193AF121228Homo sapiensthyroid hormone receptor-associated protein complex component TRAP951037100194G04072Homo sapiensHuman secreted protein, SEQ ID NO: 8153.12462195Y69384Homo sapiensArnino acid sequence of a 14274 receptor protein.119100	790		Homo sapiens	CGI-90 protein		96
AF121228	791		Homo sapiens			
complex component TRAP95 94 G04072 Homo sapiens Human secreted protein, SEQ ID NO: 8153. 124 62 95 Y69384 Homo sapiens Amino acid sequence of a 14274 receptor protein. 119 100	792			Human PRO381 protein sequence.	<u> </u>	
94 G04072 Homo sapiens Human secreted protein, SEQ ID NO: 8153. 124 62 95 Y69384 Homo sapiens Amino acid sequence of a 14274 receptor protein. 119 100	793	AF121228	Homo sapiens	complex component TRAP95	1037	100
795 Y69384 Homo sapiens Amino acid sequence of a 14274 receptor 119 100 protein.	794		Homo sapiens		124	62
	795	Y69384		Amino acid sequence of a 14274 receptor	119	100
96 W40215 Homo sapiens Human macrophage antigen. 1358 99	796	W40215	Homo sapiens		1358	99

CEC	I Accession	Species	Description	1.0	10/
SEQ ID	Accession No.	opecies	Description	Smith-	%
NO:	140.			Waterman	Identity
797	AF258340	Homo sapiens	hand coll land and land and land	Score	99
798	AF159615		hepatocellular carcinoma-associated antigen 112	1151	
799	Y59863	Homo sapiens	FGF receptor activating protein 1	461	98
		Homo sapiens	Human normal uterus tissue derived protein 26.	797	99
800	W70459	Homo sapiens	Human T1-receptor ligand III splice variant 2.	572	92
801	L00073	Homo sapiens	renin	1913	93
802	P92219	Homo sapiens	CR1 protein.	11963	97
		(human)			1
803	X15357	Homo sapiens	ANP-A receptor preprotein (AA -32 to 1029)	5199	98
804	W64473	Homo sapiens	Human secreted protein from clone EC172_1.	4018	95
805	AJ243874	Homo sapiens	oligophrenin-4	2067	100
806	G01731	Homo sapiens	Human secreted protein, SEQ ID NO: 5812.	284	100
807	Z24680	Homo sapiens	garp	1562	83
808	AF171669	Homo sapiens	glycoprotein-associated amino acid transporter	1364	90
			LAT2		1
809	W70321	Homo sapiens	Secreted protein CC198 1.	1154	96
018	W74843	Homo sapiens	Human secreted protein encoded by gene 115	855	99
		· ·	clone HOVBA03.		1
811	AF108831	Homo sapiens	K:Cl cotransporter 3	4561	100
812	AF092135	Homo sapiens	PTD014	862	100
813	AF283772	Homo sapiens	similar to Homo sapiens ribosomal protein L10	784	100
!	1	Tromo supremo	encoded by GenBank Accession Number	1,04	100
			L25899		1
814	G01563	Homo sapiens	Human secreted protein, SEQ ID NO: 5644.	330	100
815	AF051151	Homo sapiens	Toll/interleukin-1 receptor-like protein 3	3850	99
816	W95630	Homo sapiens	Homo sapiens secreted protein gene clone	358	100
010	W 93030	Figure Sapiens	gn 114 1.	338	100
817	G01082	Homo sapiens		549	100
818	AF151800		Human secreted protein, SEQ ID NO: 5163.		100
819	L00352	Homo sapiens	CGI-41 protein	1106	95
820		Homo sapiens	low density lipoprotein receptor	3980	100
	X04434	Homo sapiens	IGF-I receptor	5832	99
821	G03844	Homo sapiens	Human secreted protein, SEQ ID NO: 7925.	572	100
822	AF212220	Homo sapiens	TERA	396	48
823	Y50125	Homo sapiens	Human glycophosphatidylinositol-anchored	4897	99
004	177166888	 	protein GPI-122.		L
824	AF156778	Homo sapiens	ASB-3 protein	2675	98
825	AF096322	Homo sapiens	neuronal voltage-gated calcium channel gamma-	1105	100
			2 subunit		
826	Y07972	Homo sapiens	Human secreted protein fragment #2 encoded	1540	100
			from gene 28.		
827	AB032013	Homo sapiens	potassium channel Kv8.1	2435	95
828	Y13620	Homo sapiens	BCL9	5284	96
829	Y91474	Homo sapiens	Human secreted protein sequence encoded by	541	98
			gene 24 SEQ ID NO:147.		,
830	X54232	Homo sapiens	glypican	1625	87
831	X14830	Homo sapiens	acetylcholine receptor beta-subunit preprotein	2540	100
832	Y71262	Homo sapiens	Human chondromodulin-like protein, Zchm1.	1002	100
833	G03873	Homo sapiens	Human secreted protein, SEQ ID NO: 7954.	638	96
834	AC003030	Homo sapiens	R29828 1	1389	93
835	Y38422	Homo sapiens	Human secreted protein.	964	87
836	U41557	Caenorhabditi	glycine-rich	85	36
	}	s elegans	· ·		
837	AL121889	Homo sapiens	dJ1076E17.1 (KIAA0823 protein (continues in	998	75
			AL023803))		´-
838	AJ011415	Homo sapiens	plexin-B1/SEP receptor	1580	60
839	W80398	Homo sapiens	A secreted protein encoded by clone cw1543 3.	1105	67
840	G00862	Homo sapiens	Human secreted protein, SEQ ID NO: 4943.	255	92
841	G02650	Homo sapiens	Human secreted protein, SEQ ID NO: 4743.	644	97
842	AF036717	Homo sapiens	FGFR signalling adaptor SNT-1	2629	99
843	Y73446	Homo sapiens			
UTJ	1 13470	Tionio sapiens	Human secreted protein clone yc27_1 protein	1089	100
844	G02872	Homo sapiens	sequence SEQ ID NO:114.	257	60
845			Human secreted protein, SEQ ID NO: 6953.	357	69
846	AF151810	Homo sapiens	CGI-52 protein	1443	88
847	X83378	Homo sapiens	putative chloride channel	1620	99
047	AC004883	Homo sapiens	similar to general transcription factor 2I; similar	655	96

SEQ	Accession	Species	Description	Smith-	1%
ID	No.	οραιω	Description	Waterman	1
NO:	110.			Score	Identity
110.	 	 -	to AF038969 (PID:g2827207)	Score	
848	X99886	Homo sapiens	monocyte chemotactic protein-2	160	76
849	AC005587	Homo sapiens	similar to mouse olfactory receptor 13; similar to	963	98
",	11000550	Tromo suprems	P34984 (PID:g464305)	100	70
850	AB038237	Homo sapiens	G protein-coupled receptor C51.2	1767	100
851	AF124490	Homo sapiens	ARF GTPase-activating protein GIT1	3415	98
852	Y86217	Homo sapiens	Human secreted protein HWHGU54, SEQ ID	1189	99
***	100217	Tromo sapions	NO:132.	1107	99
853	AF224741	Homo sapiens	chloride channel protein 7	3748	99
854	X17094	Homo sapiens	furin (AA 1-794)	3550	99
855	W78245	Homo sapiens	Fragment of human secreted protein encoded by	1245	99
000	1170245	Tionio sapiciis	gene 19.	1243	99
856	R97569	Homo sapiens	Interleukin-2 receptor associated protein p43.	1926	100
857	Y41765	Homo sapiens	Human PRO1083 protein sequence.	3211	99
858	AF057306	Homo sapiens	transmembrane proteolipid		
859	AK025116	Homo sapiens	unnamed protein product	481 374	84 69
860	Y41312	Homo sapiens	Human secreted protein encoded by gene 5 clone		
800	141312	Homo sapiens	HLDRM43.	824	100
862	Y25776	Homo sapiens	Human secreted protein encoded from gene 66.	895	00
863	Y74188	Homo sapiens	Human prostate tumor EST fragment derived	96	99
1 003	1 /4166	riomo sapiens	protein #375.	96	30
864	AF167473	Homo sapiens	heme-binding protein	870	1-00
865	G02532	Homo sapiens	Human secreted protein, SEQ ID NO: 6613.	870 211	99
866	X54870	Homo sapiens	Type II integral membrane protein	1201	67
867	G00700	Homo sapiens	Human secreted protein, SEQ ID NO: 4781.		100
868	Y07894	Homo sapiens		640	99
808	107654	Homo sapiens	Human secreted protein fragment encoded from gene 43.	388	88
869	J00123	Homo sapiens	preproenkephalin (1349	05
870	Y91632	Homo sapiens	Human secreted protein sequence encoded by		95
070	1 91032	Homo sapiens	gene 25 SEQ ID NO:305.	1048	98
871	L04311	Homo sapiens	GABA-alpha receptor beta-3 subunit	237	102
872	Y29988	Homo sapiens	Human cytokine family member EF-7 protein.	960	93
873	AF161382	Homo sapiens	HSPC264		94
874	G03412	Homo sapiens		1124	99
875	Y27572	Homo sapiens	Human secreted protein, SEQ ID NO: 7493. Human secreted protein encoded by gene No. 6.	464	100
876	M15530	Homo sapiens	B-cell growth factor	573 171	96
877	W63681	Homo sapiens	Human secreted protein 1.		56
878	L27867	Rattus	neurexophilin	1652 1448	99
0,0	L27607	norvegicus	пешехорини	1448	98
879	Y10835	Homo sapiens	Amino acid sequence of a human secreted	321	100
0,,	110055	Tionio sapicais	protein.	321	100
880	W88991	Homo sapiens	Polypeptide fragment encoded by gene 144.	936	100
881	AF118670	Homo sapiens	orphan G protein-coupled receptor	1971	100
882	AF208865	Homo sapiens	EDRF		
883	Y18462	Homo sapiens	cathepsin L	528 209	100
884	Y94950	Homo sapiens	Human secreted protein clone dh1073_12 protein	348	72 100
	******		sequence SEQ ID NO:106.	J40	100 .
885	AF070661	Homo sapiens	HSPC005	404	100
886	Y04315	Homo sapiens	Human secreted protein encoded by gene 23.	385	100
887	X92744	Homo sapiens	hBD-1	375	100
888	Y22496	Homo sapiens	Human secreted protein sequence clone		
-50	1-2470	1 TOTALO SAPICALS	cn621 8.	994	94
889	Y41293	Homo sapiens	Human soluble protein ZTMPO-1.	4595	99
890	G03714	Homo sapiens	'Human secreted protein, SEQ ID NO: 7795.	147	
891	AF208856	Homo sapiens	BM-014		63
892	U29195	Homo sapiens		1012	99
893	X68149	Homo sapiens	neuronal pentraxin II	2002	98
894	Y94914		Burkitt lymphoma receptor 1	1953	100
374	1 24214	Homo sapiens	Human secreted protein clone pw337_6 protein	537	100
895	W61620	Homo conion-	sequence SEQ ID NO:34.	226	
896	W61630	Homo sapiens	Clone HNFGW06 of EGFR receptor family.	326	63
896	M24110	Homo sapiens	G0S19-2 peptide precursor	481	100
	Z68747	Homo sapiens	imogen 38	2018	99
898	AF186112	Homo sapiens	neurokinin B-like protein ZNEUROK1	619	100
899	AF225420	Homo sapiens	AD025	734	100

SEQ	Accession	Species	Description	Consists.	10/
ID	No.	Species	Description	Smith- Waterman	%
NO:	No.	1	Į	Score	Identity
900	P60657	Homo sapiens	S. C. C. C. C. C. C. C. C. C. C. C. C. C.		100
	M27288		Sequence of human lipocortin.	1835	100
901		Homo sapiens	oncostatin M	1297	99
902	W85737	Homo sapiens	Polypeptide with transmembrane domain.	749	100
903	G01349	Homo sapiens	Human secreted protein, SEQ ID NO: 5430.	650	99
904	Y00261	Homo sapiens	Human secreted protein encoded by gene 4.	1133	99
905	AF039688	Homo sapiens	antigen NY-CO-3	771	99
906	AB007836	Homo sapiens	Hic-5	2544	100
907	AB017507	Homo sapiens	Apg12	224	100
908	AK000056	Homo sapiens	unnamed protein product	1537	98
909	Y86299	Homo sapiens	Human secreted protein HFOXB55, SEQ ID	427	100
		1101110 2441111	NO:214.	1 '2'	100
910	AF231023	Homo sapiens	protocadherin Flamingo 1	7393	99
911	Y14134	Homo sapiens	Vascular endothelial cell growth inhibitor beta	1319	100
711	114154	Tionio sapiciis	protein sequence.	1313	100
912	Z90420	II		1050	100
912	2.90420	Homo sapiens	Human GDF-3 (hGDF-3) polypeptide encoding	1950	100
012	7/1005	 	cDNA.	<u> </u>	
913	Y19757	Homo sapiens	SEQ ID NO 475 from WO9922243.	1361	100
914	G03172	Homo sapiens	Human secreted protein, SEQ ID NO: 7253.	112	48
915	U14971	Homo sapiens	ribosomal protein S9	886	90
916	AF172854	Homo sapiens	cardiotrophin-like cytokine CLC	1204	99
917	AC005525	Homo sapiens	F22162 1	1963	100
918	AF166350	Homo sapiens	ST7 protein	4711	99
919	Y87285	Homo sapiens	Human signal peptide containing protein HSPP-	430	100
			62 SEQ ID NO:62.	.50	
920	Y36131	Homo sapiens	Human secreted protein #3.	465	88
921	AF193766	Homo sapiens	cytokine-like protein C17	724	100
922	Y95013	Homo sapiens	Human secreted protein vc48 1, SEQ ID NO:66.	357	100
923	X75208	Homo sapiens	protein tyrosine kinase-receptor		
923	Y96202	Homo sapiens		5256	100
		Homo sapiens	IkappaB kinase (IKK) binding protein, Y2H56.	813	98
925	AB039886	Homo sapiens	down-regulated in gastric cancer	785	78
926	G03368	Homo sapiens	Human secreted protein, SEQ ID NO: 7449.	55	50
927	Y48606	Homo sapiens	Human breast tumour-associated protein 67.	539	100
928	Y36151	Homo sapiens	Human secreted protein #23.	668	100
929	AF110399	Homo sapiens	elongation factor Ts	1666	100
930	AF210317	Homo sapiens	facilitative glucose transporter family member	2763	99
			GLUT9		
931	Y73328	Homo sapiens	HTRM clone 082843 protein sequence.	931	100
932	G01959	Homo sapiens	Human secreted protein, SEQ ID NO: 6040.	274	100
933	U47924	Homo sapiens	B-cell receptor associated protein	1469	100
934	G03827	Homo sapiens	Human secreted protein, SEQ ID NO: 7908.	529	93
935	AB039371	Homo sapiens	mitochondrial ABC transporter 3	196	63
936	X56385	Canis	rab8	1064	100
930	730363	familiaris	1406	1064	100
937	B08906		Thuman	L	
731	סטכסטם	riomo sapiens	Human secreted protein sequence encoded by	117	44
026	112600	+,,	gene 16 SEQ ID NO:63.	-	
938	M13692	Homo sapiens	alpha-1 acid glycoprotein precursor	1064	99
939	Y53886	Homo sapiens	A suppressor of cytokine signalling protein	515	42
			designated HSCOP-6.		ĺ
940	Y16630	Homo sapiens	Human Putative Adrenomedullin Receptor	1904	99
	<u> </u>		(PAR).		}
941	AC005102	Homo sapiens	small inducible cytokine subfamily A member	627	99
		1	24]	
942	M12886	Homo sapiens	T-cell receptor beta chain	1289	81
943	AF226046	Homo sapiens	GK003	1049	98
944	Y36078	Homo sapiens	Extended human secreted protein sequence, SEQ	667	100
			ID NO. 463.	007	100
945	M22877	Homo sapiens		565	100
945	W67869		cytochrome c	565	100
340	W0/809	Homo sapiens	Human secreted protein encoded by gene 63	551	93
047	31/65050		clone HHGDB72.		
947	W67859	Homo sapiens	Human secreted protein encoded by gene 53	283	100
	l	<u> </u>	clone HBMCL41.		<u></u>
948	W85726	Homo sapiens	Novel protein (Clone BG33_7).	789	100
949	AJ242015	Homo sapiens	eMDC II protein	4236	100
950	G04075	Homo sapiens	Human secreted protein, SEQ ID NO: 8156.	567	99

CCO.	Accession	Cassies	Description	1 0 14	10/
SEQ ID	No.	Species	Description	Smith-	%
NO:	190.			Waterman	Identity
951	AF110645	II	22 72 (2.1)	Score	1.00
	- I	Homo sapiens	candidate tumor suppressor p33 ING1 homolog	1314	100
952	Y36111	Homo sapiens	Extended human secreted protein sequence, SEQ	402	70
~	1.50.00		ID NO. 496.	ļ	L
953	AB012109	Homo sapiens	APC10	990	100
954	AF246221	Homo sapiens	transmembrane protein BRI	1405	100
955	AF054986	Homo sapiens	putative transmembrane GTPase	1883	100
956	W74726	Homo sapiens	Human secreted protein fg949_3.	1879	100
957	Y27096	Homo sapiens	Human viral receptor protein (ACVRP).	1581	100
958	AJ222967	Homo sapiens	cystinosin	1920	100
959	Y53052	Homo sapiens	Human secreted protein clone df202 3 protein	587	100
	ļ	•	sequence SEQ ID NO:110.	J	
960	G02694	Homo sapiens	Human secreted protein, SEQ ID NO: 6775.	283	100
961	AF151855	Homo sapiens	CGI-97 protein	1214	96
962	U26592	Homo sapiens	diabetes mellitus type I autoantigen	250	65
963	AL050306	Homo sapiens	dJ475B7.2 (novel protein)	3796	100
964	AF078859	Homo sapiens	PTD004	2089	100
965	AB020315	Homo sapiens	homologue of mouse dkk-1 gene:Acc#	1466	100
900	AB020313	Homo Sapiens	AF030433	1400	100
966	X04571	Homo sapiens	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	6590	100
967		Liomo sapiens	precursor polypeptide (AA -22 to 1185)	6580	99
968	AF146019 AF071002	Homo sapiens	hepatocellular carcinoma antigen gene 520	993	99
	1	Homo sapiens	minK-related peptide 1; MiRP1	632	100
969	AB021227	Homo sapiens	membrane-type-5 matrix metalloproteinase	3545	100
970	AF180920	Homo sapiens	cyclin L ania-6a	1579	100
971	AF105365	Homo sapiens	K-Cl cotransporter KCC4	5621	99
972	AF083248	Homo sapiens	ribosomal protein L26 homolog	739	100
973	AJ132429	Homo sapiens	hyperpolarization-activated cyclic nucleotide	6295	100
	[_i	gated cation channel hHCN4	1	1
974	W61619	Homo sapiens	Clone HTPEF86 of TM4SF superfamily.	454	100
975	AF155100	Homo sapiens	zinc finger protein NY-REN-21 antigen	2261	100
976	AF275948	Homo sapiens	ABCA1	11763	99
977	AB026891	Homo sapiens	cystine/glutamate transporter	2552	100
978	AF117657	Homo sapiens	thyroid hormone receptor-associated protein	3348	99
		,	complex component TRAP80	1 55.5	"
979	AF044201	Rattus	neural membrane protein 35; NMP35	1570	92
		norvegicus	nosiai memerano protein se, mari se	1370	"
980	AF119297	Homo sapiens	neuroendocrine-specific protein-like protein 1	1170	99
981	AF155652	Homo sapiens	potassium channel modulatory factor	1983	99
982					
702	W88499		Human stomach carcinoma clone UD10/12		
	W88499	Homo sapiens	Human stomach carcinoma clone HP10412-	1553	99
983		Homo sapiens	encoded protein.	1553	99
983	Z56281	Homo sapiens Homo sapiens	encoded protein. interferon regulatory factor 3	1553	99 98
984	Z56281 AB026125	Homo sapiens Homo sapiens Homo sapiens	encoded protein. interferon regulatory factor 3 ART-4	1553 2012 2160	99 98 100
	Z56281	Homo sapiens Homo sapiens	encoded protein. interferon regulatory factor 3 ART-4 Fragment of human secreted protein encoded by	1553	99 98
984 985	Z56281 AB026125 Y14482	Homo sapiens Homo sapiens Homo sapiens Homo sapiens	encoded protein. interferon regulatory factor 3 ART-4 Fragment of human secreted protein encoded by gene 17.	2012 2160 172	99 98 100 70
984 985 986	Z56281 AB026125 Y14482 AB023888	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	encoded protein. interferon regulatory factor 3 ART-4 Fragment of human secreted protein encoded by gene 17. b-chemokine receptor CCR4	2012 2160 172	99 98 100 70
984 985 986 987	Z56281 AB026125 Y14482 AB023888 W27291	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	encoded protein. interferon regulatory factor 3 ART-4 Fragment of human secreted protein encoded by gene 17. b-chemokine receptor CCR4 Human H1075-1 secreted protein 5' end.	1553 2012 2160 172 1895 712	99 98 100 70 100 100
984 985 986	Z56281 AB026125 Y14482 AB023888	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Manduca	encoded protein. interferon regulatory factor 3 ART-4 Fragment of human secreted protein encoded by gene 17. b-chemokine receptor CCR4	2012 2160 172	99 98 100 70
984 985 986 987 988	Z56281 AB026125 Y14482 AB023888 W27291 AF153450	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Manduca sexta	encoded protein. interferon regulatory factor 3 ART-4 Fragment of human secreted protein encoded by gene 17. b-chemokine receptor CCR4 Human H1075-1 secreted protein 5' end. juvenile hormone esterase binding protein	1553 · · · · · · · · · · · · · · · · · ·	99 98 100 70 100 100 32
984 985 986 987 988 989	Z56281 AB026125 Y14482 AB023888 W27291 AF153450 G03697	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Manduca sexta Homo sapiens	encoded protein. interferon regulatory factor 3 ART-4 Fragment of human secreted protein encoded by gene 17. b-chemokine receptor CCR4 Human H1075-1 secreted protein 5' end. juvenile hormone esterase binding protein Human secreted protein, SEQ ID NO: 7778.	1553 · · · · · · · · · · · · · · · · · ·	99 98 100 70 100 100 32
984 985 986 987 988	Z56281 AB026125 Y14482 AB023888 W27291 AF153450	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Manduca sexta	encoded protein. interferon regulatory factor 3 ART-4 Fragment of human secreted protein encoded by gene 17. b-chemokine receptor CCR4 Human H1075-1 secreted protein 5' end. juvenile hormone esterase binding protein Human secreted protein, SEQ ID NO: 7778. potassium large conductance calcium-activated	1553 · · · · · · · · · · · · · · · · · ·	99 98 100 70 100 100 32
984 985 986 987 988 989	Z56281 AB026125 Y14482 AB023888 W27291 AF153450 G03697 AF204159	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Manduca sexta Homo sapiens	encoded protein. interferon regulatory factor 3 ART-4 Fragment of human secreted protein encoded by gene 17. b-chemokine receptor CCR4 Human H1075-1 secreted protein 5' end. juvenile hormone esterase binding protein Human secreted protein, SEQ ID NO: 7778. potassium large conductance calcium-activated channel beta 3a subunit	1553 2012 2160 172 1895 712 226 194 1486	99 98 100 70 100 100 32 88 100
984 985 986 987 988 989 990 .	Z56281 AB026125 Y14482 AB023888 W27291 AF153450 G03697 AF204159 G02061	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Manduca sexta Homo sapiens Homo sapiens	encoded protein. interferon regulatory factor 3 ART-4 Fragment of human secreted protein encoded by gene 17. b-chemokine receptor CCR4 Human H1075-1 secreted protein 5' end. juvenile hormone esterase binding protein Human secreted protein, SEQ ID NO: 7778. potassium large conductance calcium-activated channel beta 3a subunit Human secreted protein, SEQ ID NO: 6142.	1553 2012 2160 172 1895 712 226 194 1486 558	99 98 100 70 100 100 32 88 100
984 985 986 987 988 989	Z56281 AB026125 Y14482 AB023888 W27291 AF153450 G03697 AF204159	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Manduca sexta Homo sapiens Homo sapiens Caenorhabditi	encoded protein. interferon regulatory factor 3 ART-4 Fragment of human secreted protein encoded by gene 17. b-chemokine receptor CCR4 Human H1075-1 secreted protein 5' end. juvenile hormone esterase binding protein Human secreted protein, SEQ ID NO: 7778. potassium large conductance calcium-activated channel beta 3a subunit	1553 2012 2160 172 1895 712 226 194 1486	99 98 100 70 100 100 32 88 100
984 985 986 987 988 989 990 991 992	Z56281 AB026125 Y14482 AB023888 W27291 AF153450 G03697 AF204159 G02061 AL031266	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Manduca sexta Homo sapiens Homo sapiens Homo sapiens Caenorhabditi s elegans	encoded protein. interferon regulatory factor 3 ART-4 Fragment of human secreted protein encoded by gene 17. b-chemokine receptor CCR4 Human H1075-1 secreted protein 5' end. juvenile hormone esterase binding protein Human secreted protein, SEQ ID NO: 7778. potassium large conductance calcium-activated channel beta 3a subunit Human secreted protein, SEQ ID NO: 6142. VM106R.1	1553 2012 2160 172 1895 712 226 194 1486 558 327	99 98 100 70 100 100 32 88 100 99 40
984 985 986 987 988 989 990 991 992	Z56281 AB026125 Y14482 AB023888 W27291 AF153450 G03697 AF204159 G02061 AL031266 Y66749	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Manduca sexta Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	encoded protein. interferon regulatory factor 3 ART-4 Fragment of human secreted protein encoded by gene 17. b-chemokine receptor CCR4 Human H1075-1 secreted protein 5' end. juvenile hormone esterase binding protein Human secreted protein, SEQ ID NO: 7778. potassium large conductance calcium-activated channel beta 3a subunit Human secreted protein, SEQ ID NO: 6142. VM106R I Membrane-bound protein PRO1124.	1553 2012 2160 172 1895 712 226 194 1486 558	99 98 100 70 100 100 32 88 100
984 985 986 987 988 989 990 991 992 993 994	Z56281 AB026125 Y14482 AB023888 W27291 AF153450 G03697 AF204159 G02061 AL031266 Y66749 G01246	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Manduca sexta Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Caenorhabditi s elegans Homo sapiens	encoded protein. interferon regulatory factor 3 ART-4 Fragment of human secreted protein encoded by gene 17. b-chemokine receptor CCR4 Human H1075-1 secreted protein 5' end. juvenile hormone esterase binding protein Human secreted protein, SEQ ID NO: 7778. potassium large conductance calcium-activated channel beta 3a subunit Human secreted protein, SEQ ID NO: 6142. VM106R.1	1553 2012 2160 172 1895 712 226 194 1486 558 327	99 98 100 70 100 100 32 88 100 99 40
984 985 986 987 988 989 990 991 992	Z56281 AB026125 Y14482 AB023888 W27291 AF153450 G03697 AF204159 G02061 AL031266 Y66749	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Manduca sexta Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	encoded protein. interferon regulatory factor 3 ART-4 Fragment of human secreted protein encoded by gene 17. b-chemokine receptor CCR4 Human H1075-1 secreted protein 5' end. juvenile hormone esterase binding protein Human secreted protein, SEQ ID NO: 7778. potassium large conductance calcium-activated channel beta 3a subunit Human secreted protein, SEQ ID NO: 6142. VM106R I Membrane-bound protein PRO1124.	1553 2012 2160 172 1895 712 226 194 1486 558 327	99 98 100 70 100 100 32 88 100 99 40
984 985 986 987 988 989 990 991 992 993 994	Z56281 AB026125 Y14482 AB023888 W27291 AF153450 G03697 AF204159 G02061 AL031266 Y66749 G01246	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Manduca sexta Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Caenorhabditi s elegans Homo sapiens	encoded protein. interferon regulatory factor 3 ART-4 Fragment of human secreted protein encoded by gene 17. b-chemokine receptor CCR4 Human H1075-1 secreted protein 5' end. juvenile hormone esterase binding protein Human secreted protein, SEQ ID NO: 7778. potassium large conductance calcium-activated channel beta 3a subunit Human secreted protein, SEQ ID NO: 6142. VM106R.1 Membrane-bound protein PRO1124. Human secreted protein, SEQ ID NO: 5327.	1553 2012 2160 172 1895 712 226 194 1486 558 327 4730 141	99 98 100 70 100 100 32 88 100 99 40
984 985 986 987 988 989 990 991 992 993 994 995	Z56281 AB026125 Y14482 AB023888 W27291 AF153450 G03697 AF204159 G02061 AL031266 Y66749 G01246 AF133845	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Manduca sexta Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Caenorhabditi s elegans Homo sapiens Homo sapiens	encoded protein. interferon regulatory factor 3 ART-4 Fragment of human secreted protein encoded by gene 17. b-chemokine receptor CCR4 Human H1075-1 secreted protein 5' end. juvenile hormone esterase binding protein Human secreted protein, SEQ ID NO: 7778. potassium large conductance calcium-activated channel beta 3a subunit Human secreted protein, SEQ ID NO: 6142. VM106R 1 Membrane-bound protein PRO1124. Human secreted protein, SEQ ID NO: 5327. corin	1553 2012 2160 172 1895 712 226 194 1486 558 327 4730 141 5811	99 98 100 70 100 100 32 88 100 99 40 99 77 99
984 985 986 987 988 989 990 991 992 993 994 995	Z56281 AB026125 Y14482 AB023888 W27291 AF153450 G03697 AF204159 G02061 AL031266 Y66749 G01246 AF133845	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Manduca sexta Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Caenorhabditi s elegans Homo sapiens Homo sapiens Homo sapiens	encoded protein. interferon regulatory factor 3 ART-4 Fragment of human secreted protein encoded by gene 17. b-chemokine receptor CCR4 Human H1075-1 secreted protein 5' end. juvenile hormone esterase binding protein Human secreted protein, SEQ ID NO: 7778. potassium large conductance calcium-activated channel beta 3a subunit Human secreted protein, SEQ ID NO: 6142. VM106R 1 Membrane-bound protein PRO1124. Human secreted protein, SEQ ID NO: 5327. corin thyroid hormone receptor-associated protein	1553 2012 2160 172 1895 712 226 194 1486 558 327 4730 141 5811 4999	99 98 100 70 100 100 32 88 100 99 40 99 77 99
984 985 986 987 988 989 990 991 992 993 994 995 996	Z56281 AB026125 Y14482 AB023888 W27291 AF153450 G03697 AF204159 G02061 AL031266 Y66749 G01246 AF133845 AF117756 W62066	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Caenorhabditi s elegans Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	encoded protein. interferon regulatory factor 3 ART-4 Fragment of human secreted protein encoded by gene 17. b-chemokine receptor CCR4 Human H1075-1 secreted protein 5' end. juvenile hormone esterase binding protein Human secreted protein, SEQ ID NO: 7778. potassium large conductance calcium-activated channel beta 3a subunit Human secreted protein, SEQ ID NO: 6142. VM106R I Membrane-bound protein PRO1124. Human secreted protein, SEQ ID NO: 5327. corin thyroid hormone receptor-associated protein complex component TRAP150 Human stern cell antigen 2.	1553 2012 2160 172 1895 712 226 194 1486 558 327 4730 141 5811 4999	99 98 100 70 100 100 32 88 100 99 40 99 100 99
984 985 986 987 988 989 990 991 992 993 994 995 996	Z56281 AB026125 Y14482 AB023888 W27291 AF153450 G03697 AF204159 G02061 AL031266 Y66749 G01246 AF133845 AF117756	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Manduca sexta Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Caenorhabditi s elegans Homo sapiens Homo sapiens Homo sapiens	encoded protein. interferon regulatory factor 3 ART-4 Fragment of human secreted protein encoded by gene 17. b-chemokine receptor CCR4 Human H1075-1 secreted protein 5' end. juvenile hormone esterase binding protein Human secreted protein, SEQ ID NO: 7778. potassium large conductance calcium-activated channel beta 3a subunit Human secreted protein, SEQ ID NO: 6142. VM106R I Membrane-bound protein PRO1124. Human secreted protein, SEQ ID NO: 5327. corin thyroid hormone receptor-associated protein complex component TRAP150 Human stem cell antigen 2. Human secreted protein sequence SEQ ID	1553 2012 2160 172 1895 712 226 194 1486 558 327 4730 141 5811 4999	99 98 100 70 100 100 32 88 100 99 40 99 77 99 100
984 985 986 987 988 989 990 991 992 993 994 995 996 997 998	Z56281 AB026125 Y14482 AB023888 W27291 AF153450 G03697 AF204159 G02061 AL031266 Y66749 G01246 AF133845 AF117756 W62066 Y87173	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	encoded protein. interferon regulatory factor 3 ART-4 Fragment of human secreted protein encoded by gene 17. b-chemokine receptor CCR4 Human H1075-1 secreted protein 5' end. juvenile hormone esterase binding protein Human secreted protein, SEQ ID NO: 7778. potassium large conductance calcium-activated channel beta 3a subunit Human secreted protein, SEQ ID NO: 6142. VM106R I Membrane-bound protein PRO1124. Human secreted protein, SEQ ID NO: 5327. corin thyroid hormone receptor-associated protein complex component TRAP150 Human stem cell antigen 2. Human secreted protein sequence SEQ ID NO: 212.	1553 2012 2160 172 1895 712 226 194 1486 558 327 4730 141 5811 4999 284 725	99 98 100 70 100 100 32 88 100 99 40 99 100 99 100 93 100
984 985 986 987 988 989 990 991 992 993 994 995 996	Z56281 AB026125 Y14482 AB023888 W27291 AF153450 G03697 AF204159 G02061 AL031266 Y66749 G01246 AF133845 AF117756 W62066	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Caenorhabditi s elegans Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	encoded protein. interferon regulatory factor 3 ART-4 Fragment of human secreted protein encoded by gene 17. b-chemokine receptor CCR4 Human H1075-1 secreted protein 5' end. juvenile hormone esterase binding protein Human secreted protein, SEQ ID NO: 7778. potassium large conductance calcium-activated channel beta 3a subunit Human secreted protein, SEQ ID NO: 6142. VM106R I Membrane-bound protein PRO1124. Human secreted protein, SEQ ID NO: 5327. corin thyroid hormone receptor-associated protein complex component TRAP150 Human stem cell antigen 2. Human secreted protein sequence SEQ ID	1553 2012 2160 172 1895 712 226 194 1486 558 327 4730 141 5811 4999	99 98 100 70 100 100 32 88 100 99 40 99 100 99

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
1002	G01234	Homo sapiens	Human secreted protein, SEQ ID NO: 5315.	398	96
1003	W73420	Homo sapiens	Human secreted protein encoded by Gene No. 24.	2150	100
1004	X12791	Homo sapiens	19kD SRP-protein (AA 1 - 144)	742	100
1005	M23323	Homo sapiens	membrane protein	642	100
1006	X63745	Homo sapiens	KDEL receptor	326	98
1007	Y35997	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 382.	824	99
1008	AB032918	Hylobates moloch	dopamine receptor D4	92	35
1009	Y91680	Homo sapiens	Human secreted protein sequence encoded by gene 81 SEQ ID NO:353.	1372	99
1010	AL136125	Homo sapiens	dJ304B14.1 (novel protein)	825	98
1011	G03733	Homo sapiens	Human secreted protein, SEQ ID NO: 7814.	379	98
1012	Y17531	Homo sapiens	Human secreted protein clone BL205 14 protein.	818	97
1013	G00724	Homo sapiens	Human secreted protein, SEQ ID NO: 4805.	462	100
1014	AF288092	Naegleria gruberi	haem lyase	114	37
1015	AB045292	Homo sapiens	M83 protein	3867	99
1016	X15940	Homo sapiens	ribosomal protein L31 (AA 1-125)	644	100
1017	Y94873	Homo sapiens	Human protein clone HP02632.	1876	100
1018	AL024498	Homo sapiens	dJ417M14.1 (novel protein)	589	100
1019	X83425	Homo sapiens	Lutheran blood group glycoprotein	3054	99
1020	W03516	Homo sapiens	Prostaglandin DP receptor.	1864	100
1021	G03960	Homo sapiens	Human secreted protein, SEQ ID NO: 8041.	398	100
1022	Y91689	Homo sapiens	Human secreted protein sequence encoded by gene 93 SEQ ID NO:362.	768	100
1023	AE000660	Homo sapiens	hADV36S1	573	100
1024	AF132965	Homo sapiens	CGI-31 protein	1550	100
1025	W92380	Homo sapiens	Human TR-interacting protein S103a.	1466	97
1026	R66278	Homo sapiens	Therapeutic polypeptide from glioblastoma cell line.	830	100
1027	X65614	Homo sapiens	S100P calcium-binding protein	476	100
1028	Y41741	Homo sapiens	Human PRO704 protein sequence.	1323	100
1029	AJ001014	Homo sapiens	RAMP1	806	100
1030	W63682	Homo sapiens	Human secreted protein 2.	1354	99
1031	AK023007	Homo sapiens	unnamed protein product	766	100
1032	W97900	Homo sapiens	Human SR-BI class B scavenger.	2672	99
1033	Y82453	Homo sapiens	Human TGC-440 sccretory protein SEQ ID NO:1.	639	99
1034	Y73473	Homo sapiens	Human secreted protein clone yd178_1 protein sequence SEQ ID NO:168.	752	93
1035	Y86468	Homo sapiens	Human gene 48-encoded protein fragment, SEQ ID NO:383.	96	90
1036	U09813	Homo sapiens	mitochondrial ATP synthase subunit 9 precursor	698	100
1037	AJ242832	Homo sapiens	calpain	3699	99
1038	X66403	Homo sapiens	acetylcholine receptor epsilon subunit CHRNE	2574	100
1039	AJ242730	Homo sapiens	polyhomeotic 2	1310	100
1040	AF169968	Mus musculus	DNA binding protein DESRT	1453	80
1041	X52563	Bos taurus	permability increasing protein	383	29
1042	G00368	Homo sapiens	Human secreted protein, SEQ ID NO: 4449.	75	50
1043	G02532	Homo sapiens	Human secreted protein, SEQ ID NO: 6613.	60	53
1044	M94582	Homo sapiens	interleukin 8 receptor B	1850	100
1045	AL080239	Homo sapiens	bG256O22.1 (similar to IGFALS (insulin-like growth factor binding protein, acid labile subunit))	1704	50
1046	AF125101	Homo sapiens	HSPC040 protein	580	100
1047	W74809	Homo sapiens	Human secreted protein encoded by gene 81 clone HMWDN32.	176	100
1048	AL022238	Homo sapiens	dJ1042K10.4 (novel protein)	2201	100
1049	W88667	Homo sapiens	Secreted protein encoded by gene 134 clone	1559	99
		·	HAIBP89.		•

SEQ	Accession	Species	Description 1	I C-:	10/
ID	No.	Species	Description	Smith-	%
NO:	140.	Į	'	Waterman	Identity
1051	W78324	Ilamaia		Score	<u> </u>
1021	W /8324	Homo sapiens	Fragment of human secreted protein encoded by	1318	98
1050	Y21851	 	gene 81.		<u> </u>
1052	Y21851	Homo sapiens	Human signal peptide-contianing protein (SIGP)	1643	95
	1	 	(clone ID 2328134).	<u> </u>	
1053	AL163815	Arabidopsis	putative protein	661	62
		thaliana			<u>l</u>
1054	Y76200	Homo sapiens	Human secreted protein encoded by gene 77.	262	100
1055	AJ276567	Homo sapiens	TC10-like Rho GTPase	1160	100
1056	Y27620	Homo sapiens	Human secreted protein encoded by gene No. 54.	154	96
1057	D14530	Homo sapiens	ribosomal protein	745	100
1058	AF132000	Homo sapiens	TADA1 protein	1132	100
1059	AL031778	Homo sapiens	dJ34B21.1 (novel BZRP (benzodiazapine	920	100
	1	1 -	receptor (peripheral) (MBR, PBR, PBKS, IBP,		-55
	1	1	Isoquinoline-binding protein)) LIKE protein)	į.	
1060	AF227135	Homo sapiens	candidate taste receptor T2R9	134	33
1061	Y27575	Homo sapiens	Human secreted protein encoded by gene No. 9.	1392	100
1062	Z11697	Homo sapiens	HB15	1088	100
1063	AF123757	Homo sapiens	putative transmembrane protein	819	100
1064	AF155135	Homo sapiens	novel retinal pigment epithelial cell protein	2932	99
1065	Y41674	Homo sapiens	Human channel-related molecule HCRM-2.		
1065				936	99
1067	AJ250042	Homo sapiens	Rab5 GDP/GTP exchange factor homologue	2575	100
1067	Y36087	Homo sapiens	Extended human secreted protein sequence, SEQ	770	85
1060	1,010.00		ID NO. 472.		<u> </u>
1068	Y94959	Homo sapicns	Human secreted protein clone mc300_1 protein	301	100
			sequence SEQ ID NO:124.	L	<u> </u>
1069	Y94959	Homo sapiens	Human secreted protein clone mc300_1 protein	301	100
			sequence SEQ ID NO:124.	<u> </u>	ł
1070	W64535	Homo sapiens	Human leukocyte cell clone HP00804 protein.	2014	99
1071	X03145	Homo sapiens	pot. ORF III	148	50
1072	AL031177	Homo sapiens	dJ889M15.3 (novel protein)	821	91
1073	X82200	Homo sapiens	gpStaf50	249	62
1074	G03213	Homo sapiens	Human secreted protein, SEQ ID NO: 7294.	99	47
1075	Y36233	Homo sapiens	Human secreted protein encoded by gene 10.	506	55
1076	G03187	Homo sapiens	Human secreted protein, SEQ ID NO: 7268.	424	98
1077	L25899	Homo sapiens	ribosomal protein L10	332	76
1078	Y91447	Homo sapiens	Human secreted protein sequence encoded by	898	97
		1	gene 48 SEQ ID NO:168.		1 "
1079	G01862	Homo sapiens	Human secreted protein, SEQ ID NO: 5943.	290	89
1080	AB039723	Homo sapiens	WNT receptor frizzled-3	1376	92
1081	AB020527	Homo sapiens	Na/PO4 cotransporter homolog	269	100
1082	L13802	Homo sapiens	ribosmal protein small subunit	499	
1083	W75098	Homo sapiens			80
1005	W 73098	rionio sapiciis	Human secreted protein encoded by gene 42 clone HSXBI25.	143	81
1084	G03564	110		00	<u> </u>
1085	G04063	Homo sapiens	Human secreted protein, SEQ ID NO: 7645.	83	51
		Homo sapiens	Human secreted protein, SEQ ID NO: 8144.	88	43
1086	AF090942	Homo sapiens	PRO0657	124	64
1087	G00517	Homo sapiens	Human secreted protein, SEQ ID NO: 4598.	129	41
1088	G04091	Homo sapiens	Human secreted protein, SEQ ID NO: 8172.	126	36
1089	AF140631	Homo sapiens	G-protein coupled receptor 14	364	82
1090	G04063	Homo sapiens	Human secreted protein, SEQ ID NO: 8144.	114	32
1091	S72304	Mus sp.	LMW G-protein	146 .	83
1092	W88708	Homo sapiens	Secreted protein encoded by gene 175 clone	405	100
	1		HEMAM41.		
1093	W85612	Homo sapiens	Secreted protein clone fh123_5.	4358	97
1094	Y53012	Homo sapiens	Human secreted protein clone pm514 4 protein	1013	99
	ł		sequence SEQ ID NO:30.	*	
1095	Y92345	Homo sapiens	Human cancer associated antigen precursor from	409	100
-]	clone NY-REN-62.	.07	100
1096	AF090942	Homo sapiens	PRO0657	147	60
1097	L24521	Homo sapiens	transformation-related protein	166	58
1098	X56932	Homo sapiens	23 kD highly basic protein	490	70
1099	G04063	Homo sapiens	Human secreted protein, SEQ ID NO: 8144.	83	
1100	Y02693	Homo sapiens	Human secreted protein encoded by gene 44		35
	102093	Tromo sapiciis	clone HTDAD22.	149	59
	<u> </u>	<u> </u>	VIUIL III DAD44.		

SEQ	Accession	Species	Description	Smith-	1%
ID `	No.	•		Waterman	Identity
NO:	<u> </u>			Score	
1101	AF119851	Homo sapiens	PRO1722	183	72
1102	G04086	Homo sapiens	Human secreted protein, SEQ ID NO: 8167.	207	62
1103	G04063	Homo sapiens	Human secreted protein, SEQ ID NO: 8144.	91	52
1104	X74856	Mus	ribosomal protein L28	128	69
		musculus		}	
1105	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	130	62
1106	G03133	Homo sapiens	Human secreted protein, SEQ ID NO: 7214.	122	48
1107	G03040	Homo sapiens	Human secreted protein, SEQ ID NO: 7121.	69	43
1108	AF039942	Homo sapiens	HCF-binding transcription factor Zhangfei	744	99
1109	AF201951	Homo sapiens	high affinity immunoglobulin epsilon receptor	738	94
			beta subunit	}	
1110	AF111108	Mus	transient receptor potential 2	223	79
	<u> </u>	musculus			1
1111	AF119900	Homo sapiens	PRO2822	144	59
1112	Y16589	Homo sapiens	A protein that interacts with presenilins.	265	39
1113	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	178	67
1114	Y02999	Homo sapiens	Fragment of human secreted protein encoded by	164	63
1115	7,00014		gene 121.		
1115	Y30811	Homo sapiens	Human secreted protein encoded from gene 1.	1217	99
1116	X51394	Xenopus	APEG precursor protein	130	40
1117	1427026	laevis			<u> </u>
	M27826	Homo sapiens	neutral protease large subunit	442	65
1118	G03371	Homo sapiens	Human secreted protein, SEQ ID NO: 7452.	72	60
1119	G03602	Homo sapiens	Human secreted protein, SEQ ID NO: 7683.	491	97
1120	Y35906	Homo sapiens	Extended human secreted protein sequence, SEQ	244	97
1121	G03714	11	ID NO. 155.		L
1121		Homo sapiens	Human secreted protein, SEQ ID NO: 7795.	122	65
1123	Y00337	Homo sapiens	Human secreted protein encoded by gene 81.	110	90
1123	AF084830 AF212862	Homo sapiens	two pore domain K+ channel; TASK-2	703	94
1124	W64469	Homo sapiens	membrane interacting protein of RGS16	442	88
1126	G01361	Homo sapiens	Human secreted protein from clone CW795_2.	191	53
1127	G01361	Homo sapiens	Human secreted protein, SEQ ID NO: 5442.	154	100
1128	Y84320	Homo sapiens	Human secreted protein, SEQ ID NO: 5442.	165	100
1120	1 64320	Homo sapiens	Human cardiovascular system associated protein kinase-1.	815	99
1129	G02105	Homo sapiens	Human secreted protein, SEQ ID NO: 6186.	88	73
1130	Y32923	Homo sapiens	Transmembrane domain containing protein clone	700	100
1150	132,23	Tionio sapicis	HP01512.	700	100
1131	Y29817	Homo sapiens	Human synapse related glycoprotein 2.	260	91
1132	Y91644	Homo sapiens	Human secreted protein sequence encoded by	525	96
		Troine saprens	gene 43 SEQ ID NO:317.		1 30
1133	Y91449	Homo sapiens	Human secreted protein sequence encoded by	542	100
			gene 49 SEQ ID NO:170.	342	100
1134	AB017908	Homo sapiens	4F2 light chain	2399	93
1135	X51760	Homo sapiens	zinc finger protein (583 AA)	312	55
1136	Y99426	Homo sapiens	Human PRO1604 (UNQ785) amino acid	917	72
	L		sequence SEQ ID NO:308.		[-
1137	G03790	Homo sapiens	Human secreted protein, SEQ ID NO: 7871.	102	50
1138	AF155106	Homo sapiens	NY-REN-36 antigen	768	91
1139	AL031055	Homo sapiens	dJ28H20.1 (novel protein similar to membrane	117	50
	L		transport proteins)		
1140	AF011359	Bos taurus	regulator of G-protein signaling 7	138	96
1141	Y70018	Homo sapiens	Human Protease and associated protein-12	623	100
			(PPRG-12).		
1142	G04091	Homo sapiens	Human secreted protein, SEQ ID NO: 8172.	113	38
1143	AB030235	Canis	D4 dopamine receptor	89	48
		familiaris			
1144	Y94922	Homo sapiens	Human secreted protein clone pv6_1 protein	539	88
		<u> </u>	sequence SEQ ID NO:50.		
1145	X99962	Homo sapiens	rab-related GTP-binding protein	398	96
1146	G03807	Homo sapiens	Human secreted protein, SEQ ID NO: 7888.	168	79
1147	G03712	Homo sapiens	Human secreted protein, SEQ ID NO: 7793.	512	85
1148	Y28279	Homo sapiens	Human G-protein coupled receptor GRIR-1.	705	76
1149	U13642	Caenorhabditi	exon 5 similar to transmembrane domain of S.	247	36

SEQ	Accession	Species	Description	Smith-	1%
ID NO:	No.	<u> </u>		Waterman Score	Identity
	000400	s elegans	cerevisiae zinc resistance protein		
1150	G03438 G01003	Homo sapiens	Human secreted protein, SEQ ID NO: 7519.	117	62
1151		Homo sapiens	Human secreted protein, SEQ ID NO: 5084.	181,	80
1152	G03798	Homo sapiens	Human secreted protein, SEQ ID NO: 7879.	198	63
1153	X88799	Oryza sativa	DNA binding protein	95	41
	D85245	Homo sapiens	TR3beta	155	96
1155	R74272	Homo sapiens	Tumour suppressor protein, p53.	341	87
1156	Y86265	Homo sapiens	Human secreted protein HUSXE77, SEQ ID NO:180.	99	41
1157·	G02577	Homo sapiens	Human secreted protein, SEQ ID NO: 6658.	263	98
1158	AF104334	Homo sapiens	putative organic anion transporter	185	42
1159	G01393	Homo sapiens	Human secreted protein, SEQ ID NO: 5474.	173	57
1160	W75771	Homo sapiens	Human GTP binding protein APD08.	224	81
1161	AF216833	Homo sapiens	M-ABC2 protein	410	83
1162	W67816	Homo sapiens	Human secreted protein encoded by gene 10 clone HCEMU42.	1156	100
1163	AF119851	Homo sapiens	PRO1722	230	70
1164	Y87252	Homo sapiens	Human signal peptide containing protein HSPP- 29 SEQ ID NO:29.	113	31
1165	W64537	Homo sapiens	Human liver cell clone HP01148 protein.	338	82
1166	AF269286	Homo sapiens	HC6	134	64
1167	Y14482	Homo sapiens	Fragment of human secreted protein encoded by gene 17.	149	51
1168	D90789	Escherichia coli	Dipeptide transport system permease protein DppC.	411	90
1169	R63783	Homo sapiens	TG0847 protein.	344	90
1170	Y45274	Homo sapiens	Human secreted protein encoded from gene 18.	478	98
1171	D64154	Homo sapiens	Mr 110,000 antigen	347	96
1172	AB026256	Homo sapiens	organic anion transporter OATP-B	311	67
1173	G00357	Homo sapiens	Human secreted protein, SEQ ID NO: 4438.	60	52
1174	D87717	Homo sapiens	similar to human GTPase-activating protein(A49869)	178	59
1175	M64716	Homo sapiens	ribosomal protein	391	78
1176	R08330	Homo sapiens	Human IL-7 receptor clone H6.	285	67
1177	L06505	Homo sapiens	ribosomal protein L12	242	72
1178	AJ251885	Homo sapiens	organic cation transporter (OCT2)	276	88-
1179	G03258	Homo sapiens	Human secreted protein, SEQ ID NO: 7339.	155	71
1180	G01207	Homo sapiens	Human secreted protein, SEQ ID NO: 5288.	282	90
1181	AF181856	Rattus	tRNA selenocysteine associated protein	249	62
	711 101050	norvegicus	tra va selenocysteme associated protein	249	02
1182	AF161524	Homo sapiens	HSPC176	138	90
1183	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	282	66
1184	Y02671	Homo sapiens	Human secreted protein encoded by gene 22	107	71
1104	1020/1	Tionio sapiens	clone HMSJW18.	107	/1
1185	G03797	Homo sapiens	Human secreted protein, SEQ ID NO: 7878.	88	10
1186	G03564	Homo sapiens	Human secreted protein, SEQ ID NO: 7878. Human secreted protein, SEQ ID NO: 7645.	118	69
1187	AB032905	Hylobates	dopamine receptor D4	96	37
1188	G00956	concolor Homo capiens	Human secreted protein CDO ID MO. 5027	202	70
1189	G03258	Homo sapiens	Human secreted protein, SEQ ID NO: 5037.	292	78
1190		Homo sapiens	Human secreted protein, SEQ ID NO: 7339.	178	79
	G03361	Homo sapiens	Human secreted protein, SEQ ID NO: 7442.	324	76
1191	AF117755	Homo sapiens	thyroid hormone receptor-associated protein complex component TRAP230	187	70
192	Y70455	Homo sapiens	Human membrane channel protein-5 (MECHP-5).	202	67
193	G03052	Homo sapiens	Human secreted protein, SEQ ID NO: 7133.	99	42
194	G02607	Homo sapiens	Human secreted protein, SEQ ID NO: 6688.	192	76
195	W29661	Homo sapiens	Homo sapiens CI542_2 clone secreted protein.	2001	98
196	Y14104	Homo sapiens	Human GABAB receptor 1d protein sequence.	239	69
197	X61972	Homo sapiens	macropain subunit iota	149	90
198	G00534	Homo sapiens	Human secreted protein, SEQ ID NO: 4615.	145	51
199	Y86260	Homo sapiens	Human secreted protein HELHN47, SEQ ID NO:175.	1089	89
	G02607	Homo sapiens	Human secreted protein, SEQ ID NO: 6688.	154	57

SEQ	Accession	Species	Description	Smith-	1%
ID	No.	Species	Description	Waterman	Identity
NO:	140.			Score	Identity
1201	G00838	Tioms and an	Th		50
		Homo sapiens	Human secreted protein, SEQ ID NO: 4919.	404	49
1202	M27826	Homo sapiens	neutral protease large subunit	202	1
1203	Y73424	Homo sapiens	Human secreted protein clone yi4_1 protein sequence SEQ ID NO:70.	265	61
1204	AF264014	Homo sapiens	scavenger receptor cysteine-rich type 1 protein M160 precursor	625	98
1205	Y36203	Homo sapiens	Human secreted protein #75.	219	59
1206	U78111	Gallus gallus	AQ	205	57
1207	AF095448	Homo sapiens	putative G protein-coupled receptor	416	76
1208	AF116715	Homo sapiens	PRO2829	127	75
1209	AF099137	Homo sapiens	MaxiK channel beta 2 subunit	475	95
1210	AF205718	Homo sapiens	hepatocellular carcinoma-related putative tumor suppressor	423	79
1211	Y27868	Homo sapiens	Human secreted protein encoded by gene No. 107.	224	70
1212	G00719	Homo sapiens	Human secreted protein, SEQ ID NO: 4800.	117	44
1213	G01009	Homo sapiens	Human secreted protein, SEQ ID NO: 4800. Human secreted protein, SEQ ID NO: 5090.	351	73
1213	1				
	AF090942	Homo sapiens	PRO0657	124	70
1215	Y14427	Homo sapiens	Human secreted protein encoded by gene 17 clone HSIEA14.	99	77
1216	G03905	Homo sapiens	Human secreted protein, SEQ ID NO: 7986.	173	57
1217	Y57897	Homo sapiens	Human transmembrane protein HTMPN-21.	1173	100
1218	J00194	Homo sapiens	hla-dr antigen alpha chain	454	78
1219	Y59709	Homo sapiens	Secreted protein 76-28-3-A12-FL1.	470	92
1220	W81576	Homo sapiens	EBV-induced G-protein coupled receptor (EBI-	725	100
1221	W96745	Homo sapiens	polypeptide. High affinity immunoglobulin E receptor-like	650	98
1222	Y35911	Homo sapiens	protein (IGERB). Extended human secreted protein sequence, SEQ	135	31
1223	Y00278		ID NO. 160.	260	L
		Homo sapiens	Human secreted protein encoded by gene 21.		95
1224	AF161422	Homo sapiens	HSPC304	568	90
	U14970	Homo sapiens	ribosomal protein \$5	202	95
1226	G01733	Homo sapiens	Human secreted protein, SEQ ID NO: 5814.	610	100
1227	AF099973	Mus musculus	schlafen2	333	56
1228	G01218	Homo sapiens	Human secreted protein, SEQ ID NO: 5299.	155	81
1229	AF217188	Mus musculus	YIPIB	801	63
1230	AF176813	Homo sapiens	soluble adenylyl cyclase	275	100
1231	X98333	Homo sapiens	organic cation transporter	1704	100
1232	W74955	Homo sapiens	Human secreted protein encoded by gene 77 clone HOEAS24.	212	53
1233	Y94940	Homo sapiens	Human secreted protein clone yi62_1 protein sequence SEQ ID NO:86.	526	100
1234	U76618	Mus musculus	N-RAP	482	82
1235	AF044924	Homo sapiens	hook2 protein	380	97
1236	G01459	Homo sapiens	Human secreted protein, SEQ ID NO: 5540.	417	100
1237	AF000018	Homo sapiens	adapter protein	164	84
1238	W88633	Homo sapiens	Secreted protein encoded by gene 100 clone HE8EU04.	250	90
1239	W29660	Homo sapiens	Homo sapiens CH27_1 clone secreted protein.	697	98
1240	AF004161	Oryctolagus cuniculus	peroxisomal Ca-dependent solute carrier	154	52
1241	Y92710	Homo sapiens	Human membrane associated wastein 7-1-24	709	97
1241	Y95002		Human membrane-associated protein Zsig24.		88
1242	Y44905	Homo sapiens Homo sapiens	Human secreted protein vc34_1, SEQ ID NO:44. Human potassium channel molecule ERG-LP2	908 325	100
1044	1 1001/5-	 -,.	partial protein.		
1244 1245	AF284422 Y53629	Homo sapiens Homo sapiens	cation-chloride cotransporter-interacting protein A bone marrow secreted protein designated	511 1888	97
			BMS115.		
1246	AB039371	Homo sapiens	mitochondrial ABC transporter 3	389	97
1247	Y35911	Homo sapiens	Extended human secreted protein sequence, SEQ	168	39

	Accession	Species	Description	Smith-	%
ID	No.			Waterman	Identity
NO:				Score	
			ID NO. 160.	 	
1248	AF072509	Rattus	glutamate receptor interacting protein 2	559	90
		norvegicus		}	
1249	AF247042	Homo sapiens	tandem pore domain potassium channel TRAAK	661	98
1250	B08974	Homo sapiens	Human secreted protein sequence encoded by	1087	97
	<u> </u>	_	gene 27 SEQ ID NO:131.	1	
1251	L15313	Caenorhabditi	putative	858	59
		s elegans	1	İ	ļ
1252	Y29338	Homo sapiens	Human secreted protein clone it217_2 alternate	278	75
			reading frame protein.		İ
1253	W01730	Homo sapiens	Human G-protein receptor HPRAJ70.	211	92
1254	G03074	Homo sapiens	Human secreted protein, SEQ ID NO: 7155.	294	83
1255	G01818	Homo sapiens	Human secreted protein, SEQ ID NO: 5899.	253	91
1256	AF286368	Homo sapiens	eppin-1	222	54
1257	AF220264	Homo sapiens	MOST-I	87	93
1258	G02227	Homo sapiens	Human secreted protein, SEQ ID NO: 6308.	281	78
1259	Y07970	Homo sapiens	Human secreted protein fragment #2 encoded	81	94
•]	Ţ .	from gene 26.	1	1
1260	R95332	Homo sapiens	Tumor necrosis factor receptor 1 death domain	986	100
	ì	1	ligand (clone 3TW).		1
1261	AF140674	Homo sapiens	zinc metalloprotease ADAMTS6	172	36
1262	U28369	Homo sapiens	semaphorin V	237	67
1263	Y07049	Homo sapiens	Renal cancer associated antigen precursor	288	71
			sequence.	200	1 ′′
1264	Y36153	Homo sapiens	Human secreted protein #25.	187	80
1265	Y78114	Homo sapiens	Human cytokine signal regulator CKSR-2 SEQ	723	93
			ID NO:2.	1 /23	/3
1266	Y13397	Homo sapiens	Amino acid sequence of protein PRO334.	191	100
1267	AF030558	Rattus	phosphatidylinositol 5-phosphate 4-kinase	859	95
		norvegicus	gamma	1 037	23
1268	U73167	Homo sapiens	candidate tumor suppressor gene LUCA-1	159	96
1269	AF190664	Mus	LMBR2	552	76
	1	musculus		332	10
1270	AL050332	Homo sapiens	dJ570F3.1 (homolog of the rat synaptic ras	820	98
			GTPase-activating protein p135 SynGAP)	320	'
1271	G02126	Homo sapiens	Human secreted protein, SEQ ID NO: 6207.	131	95
1272	AF125533	Homo sapiens	NADIT-cytochrome b3 reductase isoform	253	92
1273	AL035661	Homo sapiens	dJ568C11.3 (novel AMP-binding enzyme	1280	100
			similar to acetyl-coenzyme A synthethase	1200	100
		1 :		(ĺ
1274		1	I (acetate-coa ligasei)	1	1
14/4	AF064748	Mus	(acetate-coA ligase)) S3-12	3523	61
12/4	AF064748	Mus musculus	S3-12	3523	61
1274	AF064748 D17554	musculus	S3-12		
		musculus Homo sapiens	S3-12 TAXREB107	377	78
1275	D17554	musculus	S3-12 TAXREB107 Amino acid sequence of a human secreted		
1275	D17554	musculus Homo sapiens Homo sapiens	S3-12 TAXREB107 Amino acid sequence of a human secreted protein.	377 643	78 90
1275 1276	D17554 Y30715 AF146760 Y05069	musculus Homo sapiens Homo sapiens Homo sapiens	S3-12 TAXREB107 Amino acid sequence of a human secreted protein. septin 2-like cell division control protein	377 643	78 90
1275 1276 1277 1278	D17554 Y30715 AF146760 Y05069	musculus Homo sapiens Homo sapiens Homo sapiens Homo sapiens	S3-12 TAXREB107 Amino acid sequence of a human secreted protein. septin 2-like cell division control protein Human PIGR-2 protein sequence.	377 643 707 281	78 90 100 46
1275 1276 1277	D17554 Y30715 AF146760	musculus Homo sapiens Homo sapiens Homo sapiens Homo sapiens Oryctolagus	S3-12 TAXREB107 Amino acid sequence of a human secreted protein. septin 2-like cell division control protein	377 643	78 90
1275 1276 1277 1278	D17554 Y30715 AF146760 Y05069 X59668	musculus Homo sapiens Homo sapiens Homo sapiens Oryctolagus cuniculus	S3-12 TAXREB107 Amino acid sequence of a human secreted protein. septin 2-like cell division control protein Human PIGR-2 protein sequence. aorta CNG channel (rACNG)	377 643 707 281 267	78 90 100 46 85
1275 1276 1277 1278 1279	D17554 Y30715 AF146760 Y05069 X59668 G01051	musculus Homo sapiens Homo sapiens Homo sapiens Homo sapiens Oryctolagus cuniculus Homo sapiens	S3-12 TAXREB107 Amino acid sequence of a human secreted protein. septin 2-like cell division control protein Human PIGR-2 protein sequence. aorta CNG channel (rACNG) Human secreted protein, SEQ ID NO: 5132.	377 643 707 281 267	78 90 100 46 85
1275 1276 1277 1278 1279	D17554 Y30715 AF146760 Y05069 X59668 G01051 G03411	musculus Homo sapiens Homo sapiens Homo sapiens Oryctolagus cuniculus Homo sapiens Homo sapiens Homo sapiens	TAXREB107 Amino acid sequence of a human secreted protein. septin 2-like cell division control protein Human PIGR-2 protein sequence. aorta CNG channel (rACNG) Human secreted protein, SEQ ID NO: 5132. Human secreted protein, SEQ ID NO: 7492.	377 643 707 281 267 489	78 90 100 46 85 98 43
1275 1276 1277 1278 1279 1280 1281 1282	D17554 Y30715 AF146760 Y05069 X59668 G01051 G03411 AF055084	musculus Homo sapiens Homo sapiens Homo sapiens Oryctolagus cuniculus Homo sapiens Homo sapiens Homo sapiens Homo sapiens	TAXREB107 Amino acid sequence of a human secreted protein. septin 2-like cell division control protein Human PIGR-2 protein sequence. aorta CNG channel (rACNG) Human secreted protein, SEQ ID NO: 5132. Human secreted protein, SEQ ID NO: 7492, very large G-protein coupled receptor-1	377 643 707 281 267 489 120 1635	78 90 100 46 85 98 43 100
1275 1276 1277 1278 1279 1280	D17554 Y30715 AF146760 Y05069 X59668 G01051 G03411	musculus Homo sapiens Homo sapiens Homo sapiens Oryctolagus cuniculus Homo sapiens Homo sapiens Homo sapiens Homo sapiens	TAXREB107 Amino acid sequence of a human secreted protein. septin 2-like cell division control protein Human PIGR-2 protein sequence. aorta CNG channel (rACNG) Human secreted protein, SEQ ID NO: 5132. Human secreted protein, SEQ ID NO: 7492.	377 643 707 281 267 489	78 90 100 46 85 98 43
1275 1276 1277 1278 1279 1280 1281 1282 1283	D17554 Y30715 AF146760 Y05069 X59668 G01051 G03411 AF055084 AF117814	musculus Homo sapiens Homo sapiens Homo sapiens Oryctolagus cuniculus Homo sapiens Homo sapiens Homo sapiens Mus musculus	TAXREB107 Amino acid sequence of a human secreted protein. septin 2-like cell division control protein Human PIGR-2 protein sequence. aorta CNG channel (rACNG) Human secreted protein, SEQ ID NO: 5132. Human secreted protein, SEQ ID NO: 7492. very large G-protein coupled receptor-1 odd-skipped related 1 protein	377 643 707 281 267 489 120 1635 357	78 90 100 46 85 98 43 100 98
1275 1276 1277 1278 1279 1280 1281 1282	D17554 Y30715 AF146760 Y05069 X59668 G01051 G03411 AF055084	musculus Homo sapiens Homo sapiens Homo sapiens Oryctolagus cuniculus Homo sapiens Homo sapiens Homo sapiens Mus musculus Xenopus	TAXREB107 Amino acid sequence of a human secreted protein. septin 2-like cell division control protein Human PIGR-2 protein sequence. aorta CNG channel (rACNG) Human secreted protein, SEQ ID NO: 5132. Human secreted protein, SEQ ID NO: 7492, very large G-protein coupled receptor-1	377 643 707 281 267 489 120 1635	78 90 100 46 85 98 43 100
1275 1276 1277 1278 1279 1280 1281 1282 1283	D17554 Y30715 AF146760 Y05069 X59668 G01051 G03411 AF055084 AF117814 U87318	musculus Homo sapiens Homo sapiens Homo sapiens Oryctolagus cuniculus Homo sapiens Homo sapiens Mus musculus Xenopus laevis	TAXREB107 Amino acid sequence of a human secreted protein. septin 2-like cell division control protein Human PIGR-2 protein sequence. aorta CNG channel (rACNG) Human secreted protein, SEQ ID NO: 5132. Human secreted protein, SEQ ID NO: 7492. very large G-protein coupled receptor-1 odd-skipped related 1 protein NaDC-2	377 643 707 281 267 489 120 1635 357	78 90 100 46 85 98 43 100 98
1275 1276 1277 1278 1279 1280 1281 1282 1283	D17554 Y30715 AF146760 Y05069 X59668 G01051 G03411 AF055084 AF117814	musculus Homo sapiens Homo sapiens Homo sapiens Oryctolagus cuniculus Homo sapiens Homo sapiens Homo sapiens Mus musculus Xenopus laevis Mus	TAXREB107 Amino acid sequence of a human secreted protein. septin 2-like cell division control protein Human PIGR-2 protein sequence. aorta CNG channel (rACNG) Human secreted protein, SEQ ID NO: 5132. Human secreted protein, SEQ ID NO: 7492. very large G-protein coupled receptor-1 odd-skipped related 1 protein	377 643 707 281 267 489 120 1635 357	78 90 100 46 85 98 43 100 98
1275 1276 1277 1278 1279 1280 1281 1282 1283 1284	D17554 Y30715 AF146760 Y05069 X59668 G01051 G03411 AF055084 AF117814 U87318	musculus Homo sapiens Homo sapiens Homo sapiens Homo sapiens Oryctolagus cuniculus Homo sapiens Homo sapiens Homo sapiens Mus musculus Xenopus laevis Mus musculus	TAXREB107 Amino acid sequence of a human secreted protein. septin 2-like cell division control protein Human PIGR-2 protein sequence. aorta CNG channel (rACNG) Human secreted protein, SEQ ID NO: 5132. Human secreted protein, SEQ ID NO: 7492. very large G-protein coupled receptor-1 odd-skipped related 1 protein NaDC-2 Edp1 protein	377 643 707 281 267 489 120 1635 357 535	78 90 100 46 85 98 43 100 98 60
1275 1276 1277 1278 1279 1280 1281 1282 1283	D17554 Y30715 AF146760 Y05069 X59668 G01051 G03411 AF055084 AF117814 U87318	musculus Homo sapiens Homo sapiens Homo sapiens Homo sapiens Oryctolagus cuniculus Homo sapiens Homo sapiens Homo sapiens Mus musculus Xenopus laevis Mus musculus Mus musculus Mus	TAXREB107 Amino acid sequence of a human secreted protein. septin 2-like cell division control protein Human PIGR-2 protein sequence. aorta CNG channel (rACNG) Human secreted protein, SEQ ID NO: 5132. Human secreted protein, SEQ ID NO: 7492. very large G-protein coupled receptor-1 odd-skipped related 1 protein NaDC-2	377 643 707 281 267 489 120 1635 357	78 90 100 46 85 98 43 100 98
1275 1276 1277 1278 1279 1280 1281 1282 1283 1284 1285	D17554 Y30715 AF146760 Y05069 X59668 G01051 G03411 AF055084 AF117814 U87318 AF061346 AB030182	musculus Homo sapiens Homo sapiens Homo sapiens Homo sapiens Oryctolagus cuniculus Homo sapiens Homo sapiens Homo sapiens Mus musculus Xenopus laevis Mus musculus Mus musculus Mus musculus Mus musculus	TAXREB107 Amino acid sequence of a human secreted protein. septin 2-like cell division control protein Human PIGR-2 protein sequence. aorta CNG channel (rACNG) Human secreted protein, SEQ ID NO: 5132. Human secreted protein, SEQ ID NO: 7492. very large G-protein coupled receptor-1 odd-skipped related 1 protein NaDC-2 Edp1 protein contains transmembrane (TM) region	377 643 707 281 267 489 120 1635 357 535 452	78 90 100 46 85 98 43 100 98 60 68
1275 1276 1277 1278 1279 1280 1281 1282 1283 1284	D17554 Y30715 AF146760 Y05069 X59668 G01051 G03411 AF055084 AF117814 U87318	musculus Homo sapiens Homo sapiens Homo sapiens Homo sapiens Oryctolagus cuniculus Homo sapiens Homo sapiens Homo sapiens Mus musculus Xenopus laevis Mus musculus Mus musculus Synthetic	TAXREB107 Amino acid sequence of a human secreted protein. septin 2-like cell division control protein Human PIGR-2 protein sequence. aorta CNG channel (rACNG) Human secreted protein, SEQ ID NO: 5132. Human secreted protein, SEQ ID NO: 7492. very large G-protein coupled receptor-1 odd-skipped related 1 protein NaDC-2 Edp1 protein	377 643 707 281 267 489 120 1635 357 535	78 90 100 46 85 98 43 100 98 60
1275 1276 1277 1278 1279 1280 1281 1282 1283 1284 1285 1286	D17554 Y30715 AF146760 Y05069 X59668 G01051 G03411 AF055084 AF117814 U87318 AF061346 AB030182	musculus Homo sapiens Homo sapiens Homo sapiens Homo sapiens Oryctolagus cuniculus Homo sapiens Homo sapiens Homo sapiens Mus musculus Xenopus laevis Mus musculus Mus musculus Synthetic construct	TAXREB107 Amino acid sequence of a human secreted protein. septin 2-like cell division control protein Human PIGR-2 protein sequence. aorta CNG channel (rACNG) Human secreted protein, SEQ ID NO: 5132. Human secreted protein, SEQ ID NO: 7492, very large G-protein coupled receptor-1 odd-skipped related 1 protein NaDC-2 Edp1 protein contains transmembrane (TM) region immunosuppresive protein PP15	377 643 707 281 267 489 120 1635 357 535 452 582	78 90 100 46 85 98 43 100 98 60 68
1275 1276 1277 1278 1279 1280 1281 1282 1283 1284 1285	D17554 Y30715 AF146760 Y05069 X59668 G01051 G03411 AF055084 AF117814 U87318 AF061346 AB030182	musculus Homo sapiens Homo sapiens Homo sapiens Homo sapiens Oryctolagus cuniculus Homo sapiens Homo sapiens Homo sapiens Mus musculus Xenopus laevis Mus musculus Mus musculus Synthetic	TAXREB107 Amino acid sequence of a human secreted protein. septin 2-like cell division control protein Human PIGR-2 protein sequence. aorta CNG channel (rACNG) Human secreted protein, SEQ ID NO: 5132. Human secreted protein, SEQ ID NO: 7492. very large G-protein coupled receptor-1 odd-skipped related 1 protein NaDC-2 Edp1 protein contains transmembrane (TM) region	377 643 707 281 267 489 120 1635 357 535 452	78 90 100 46 85 98 43 100 98 60 68

SEQ	Accession	Species	Description	Smith-	%
ID `	No.	1		Waterman	Identity
NO:	1	1		Score	1
1290	AF038563	Homo sapiens	membrane associated guanylate kinase 2	523	100
1291	AF034837	Homo sapiens	double-stranded RNA specific adenosine	468	100
			deaminase		
1292	M15888	Bos taurus	endozepine-related protein precursor	937	87
1293	AB010692	Arabidopsis	ATP-dependent RNA helicase-like protein	636	45
		thaliana			<u> </u>
1294	AF209923	Homo sapiens	orphan G-protein coupled receptor	1570	100
1295	W67828	Homo sapiens	Human secreted protein encoded by gene 22	504	98
1296	AC004832	 	clone HFEAF41.	1 213	<u> </u>
		Homo sapiens	similar to 45 kDa secretory protein; similar to CAA10644.1 (PID:g4164418)	648	65
1297	X80035	Oryctolagus cuniculus	cysteine rich hair keratin associated protein	575	70
1298	G02645	Homo sapiens	Human secreted protein, SEQ ID NO: 6726.	223	97
1299	Y59440	Homo sapiens	Human delta3 fragment #4.	122	32
1300	W70504	Homo sapiens	Leukocyte seven times membrane-penetrating type receptor protein JEG18.	459	81
1301	Y67315	Homo sapiens	Human secreted protein BL89_13 amino acid sequence.	3916	99
1302	M77693	Homo sapiens	spermidine/spermine N1-acetyltransferase	174	96
1303	G01331	Homo sapiens	Human secreted protein, SEQ ID NO: 5412.	254	69
1304	G01491	Homo sapiens	Human secreted protein, SEQ ID NO: 5572.	747	99
1305	AF148509	Homo sapiens	alpha 1,2-mannosidase	602	98
1306	G01658	Homo sapiens	Human secreted protein, SEQ ID NO: 5739.	333	98
1307	Y90899	Homo sapiens	D1-like dopamine receptor activity modifying	332	98
	1	1	protein SEQ ID NO:1.	1	1
1308	AF033120	Homo sapiens	p53 regulated PA26-T2 nuclear protein	348	52
1309	Y73388	Homo sapiens	HTRM clone 3376404 protein sequence.	147	66
1310	AF063243	Bos taurus	ribosomal protein L30	296	90
1311	AF224494	Mus musculus	arsenite inducible RNA associated protein	688	70
1312	Y73342	Homo sapiens	HTRM clone 2709055 protein sequence.	1154	100
1313	Y99419	Homo sapiens	Human PRO1780 (UNQ842) amino acid sequence SEQ ID NO:282.	1145	78
1314	AF116667	Homo sapiens	PRO1777	433	97
1315	W75100	Homo sapiens	Human secreted protein encoded by gene 44 clone HE8CJ26.	807	97
1316	AJ272078	Homo sapiens	APOBEC-1 stimulating protein	789	100
1317	AB041533	Homo sapiens	sperm antigen	2607	98
1318	U19617	Mus	ÉIf-1	806	92
1210	1100 500	musculus			
1319	U82598	Escherichia coli	ferric enterobactin transport protein	768	100
1320	D90892	Escherichia	SORBITOL-6-PHOSPHATE 2-	709	100
	ļ	coli	DEHYDROGENASE (EC 1.1.1.140)	1	
		1	(GLUCITOL-6- PHOSPHATE	İ	
			DEHYDROGENASE) (KETOSEPHOSPHATE		
1221	17767047		REDUCTASE).		
1321	W67847	Homo sapiens	Human secreted protein encoded by gene 41	601	92
1322	AJ276101	Home series	clone HPBCJ74. GPRC5B protein	166	02
1323	AJ276101 AJ276101	Homo sapiens Homo sapiens	GPRC5B protein GPRC5B protein	466 504	93 97
1324	Y58628	Homo sapiens	Protein regulating gene expression PRGE-21.	1584	100
1325	U91561	Rattus	pyridoxine 5'-phosphate oxidase	1277	89
	07.301	norvegicus	PJ1140Anic 3 -phospitate Oxidase	12//	07
1326	AF125533	Homo sapiens	NADH-cytochrome b5 reductase isoform	1606	100
1327	Y32206	Homo sapiens	Human receptor molecule (REC) encoded by	1531	90
	}	1 10mo supions	Incyte clone 2825826.	1331	70
1328	AF151048	Homo sapiens	HSPC214	657	85
1329	Y10530	Homo sapiens	olfactory receptor	1645	100
1330	AF180681	Homo sapiens	guanine nucleotide exchange factor	4314	99
1331	AF111856	Homo sapiens	sodium dependent phosphate transporter isoform	3591	99
			NaPi-3b	1	1
1332	Y13583	Homo sapiens	G-protein coupled receptor	2171	100
1332			SURF-4		

SEQ	Accession	Species	Description	Smith-	%
ID NO: _	No.			Waterman Score	Identity
1334	Y25755	Homo sapiens	Human secreted protein encoded from gene 45.	1380	96
1335	AF152325	Homo sapiens	protocadherin gamma A5	4742	99
1336	X74070	Homo sapiens	transcription factor BTF3	639	81
1337	AF095927	Rattus norvegicus	protein phosphatase 2C	1931	95
1338	G03877	Homo sapiens	Human secreted protein, SEQ ID NO: 7958.	621	100
1339	AL008582	Homo sapiens	bK223H9.2 (ortholog of A. thaliana F23F1.8)	626	100
1340	X61615	Homo sapiens	leukemia inhibitory factor receptor	5820	99
1341	Y01519	Homo sapiens	A carcinogenesis-inhibiting protein.	7528	97
1342	AF207600	Homo sapiens	ethanolamine kinase	2372	100
1343	U54807	Rattus norvegicus	GTP-binding protein	1167	97
1344	AC020579	Arabidopsis thaliana	putative phosphoribosylformylglycinamidine synthase; 25509-29950	3283	51
1345	Y28576	Homo sapiens	Secreted peptide clone pe503_1.	944	100
1346	W74787	Homo sapiens	Human secreted protein encoded by gene 58 clone HHFHN61.	1171	100
1347	M55542	Homo sapiens	guanylate binding protein isoform I	2636	87
1348	AF183428	Homo sapiens	28.4 kDa protein	1329	100
1349	U70669	Homo sapiens	Fas-ligand associated factor 3	167	24
1350	AF295530	Homo sapiens	cardiac voltage gated potassium channel modulatory subunit	562	99

TABLE 3

1	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
1	NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
ĺ	nucl-	peptide	İ	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
1	eotide	seq-	!	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
l	seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
ı	uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
ı			1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1					residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			1		peptide		/=possible nucleotide deletion, \=possible
Ĺ			Í		sequence		nucleotide insertion
I	1	1351	Α	2	337	1	TPSLIHQAPTPCPAGLWG/PPNGHYHGS*PGC
7		1					HWPQAPHRA***GLLPPRWLGHGLPGGPAAP
1			•			İ	WAASQWVDGVAGRLPGPAWSWHASGAAPA
L							QPGPL*LLVPGSSGLPDPRDP
ſ	2	1352	A	27	100	366	IRNSSIRPMKERETKLSAKHMITCSASYDIRGL
ı			1				QIETT\YHHTPIRMAKIQKT/GHHQC**ECGAT
L							GTLIHGWWGCKVVEPLGKTVWQIPK
ſ	3	1353	A	40	3	314	HASAHASVVLKDNSELEQQLGATGAYRARA
ı							LELEAEVAEMRQMLQLEHPFVNGADKLRPD
1				ľ			SMYVHLNEL*QSLVENMLLTVVDTH\RTPI*R
1							SCNYTLALILFL
ſ	4	1354	A	74	2	292	TASALFSCPDGGSLAGFAGRRASFHLECLKR
ļ	ļ						QKDRGGDISQKTVLPLHLVHHQVAHTFGQAT
١	ì						VTCQQARQSPG*RTNPE/ALQWVLPVSDGWH
L				}			VLPLP
Γ	5	1355	Α	78	114	850	ENCRVASNLPGVFFSEDTAQSGSYMRISAHPP
١							NAGGEVSNGPKRKLTLMLNFSLPSSGLNAGA
	ſ				į		FYALSTLLNRMVIWHYPGEEVNAGRIGLTIVI
	ì				ĺ		AGMLGAVISGIWLDRSKTYKETTLVVYIMDT
ĺ	ĺ						GGAWWCYTFYLGTGDTCG*CFITAG\TMGFF
l	1		' '		ļ		MTGYLPLGFEFAVEL\SYPESEGISSGLLNISA
ļ	Ì		,]				QVFGIIFTISQGQIIDNYGTKPGNIFLCVFLTLG
1	. }		- 1	1			AALTAFIKADLRROKANKETLEN
Г	6	1356	A	81	97	376	EWFSYMLGSNMSVYHSP*SLEPLCKVLSES*A
١	ļ			1	ļ		YLRVPFIRILLNAR*IRKAYKRMSLEIKLLI/RE
l	_			1			*CLFQEMGLSLQWLYSARGDFFRATSRL
r	7	1357	A	93	2	872	TLSSACLIGDAWKELTIVAGAVSNQLLVWYP
							ATALADNKPVAPDRRISGHVGIIFSMSYLESK
l					j		GLLATASEDRSVRIWKGGDLRVPGGRVQNIG
				1			HCFGHSARVWQVKLLENYLISAGEDCVCLV
٠.							

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutarnic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutarnine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						WSHEGEILQAFRGHQGRGIRAIAAHERQAWV ITGGDDSGIRLWHLVGRGYRGLG/DLGSLLQ VP**ARYTQGCDSGWLLATAGSD*YRGPVSL *RRGQVLGAAARG*TFPVLLPAGGSSWSRGL RIVCYGQWGRSCQGCPHQHSNCCCGPDPVS WEGAQLELGPAWL
8	1358	A	106	3	350	FSSLLSGRISTLRDETGAILIDGDPAACAPIIKF LLTEELHLRGVSIYVLRHEAQIYGITPL\VCAL LI/CRRL*SDSCMRAALNDRGLYQVLILDGLV QCLGFVDSDSRKMVSTLT
9	1359	A	115	49	186	QAWAIFKGKYKEGDTGGPAVWKTRLRCALN KSSEFNEGPERERMDV
10	1360	A	123	2	1249	KGCRTQEKVDRTEVIRTCINPVYSKLFTVDFY FEEVQRLRFEVHDISSNHNGLKEADFLGGME CTLGQIVSQRKLSKSLLKHGNTAGKSSITVIA EELSGNDDYVELAFNARKLDDKDFFSKSDPF LEIFRMNDDATQQLVHRTEVVMNNLSPAWK SFKVSVNSLCSGDPDRRLKCIVWDWDSNGK HDFIGEFTSTFKEMRGAMEGKQVQWECINPK YKAKKKNYKNSGTVILNLCKHKMHSFLDYI MGGCQIQFTVAIDFTASNGDPRNSCSLHYIHP YQPNEYLKALVAVGEICQDYDSDKMFPAFGF GARIPPEYTDSHDFAINFNEDNPECAGIQGVV EAYQSCF\PKAPTFTGPTNICPHSSRKVAKFRR SEGN*HQGRAFAIIFILVDPGQVGVYSQDMGP DNPGGHFV
11	1361	A	147	614	9	ACARKQLLGRTVFIWFVGQLLGGELKGYSKT NTTSSRPASSRG\TLSSSSSSSSSLTKDALPSSL KSDSTTITSGLVFPFRSLCVNPAKSSVSESVSSI KILLSSSVKYLE*KRTSCCFPDSSESKLSQLSS DERVSMGTSSRKPTNSSSSLGALKMSATS*G SGSESPTPFFLTGLQSPPSTRPREPGLTTARNS TTLTRDC
12	1362	A	177	12	416	LIPSEPALDSLVDPRVRSRKQPFVIYPVYDTAI DTKIHFSLLDGNVGEPDMSAGFCPNHKAAM VLFLDRVYGIEVQDFLLHLLEGGFLPDLRAA ASLDT/AEIGAMDFLLS*LFTLCLMMFFFIYPFI NLLTMNVY
13	1363	A	249	535	105	WTFHRHLSPAPLIVCDQGTCVVSYYPQNIVQ MPDTQMEQGLN/HLFLDGNA*PHSVECYCPS TFEIAIKITSFVLYFHRYRAPEVLLRSSVYSSPI DVWAVGSIMAELYMLRPLFPGTSEVDEIFKIC QVLGTPKKVSTLVPKLL
14	1364	A	254	572	201	YLLTXIGNLMMLLVINADSCLRTXM*FFLGH FFFLDICYSSVTAQDAAEFPVS*KPILVWGYIT *SFFFIFSWGTNGCLLSAITYACYAAICHPLLS TMVMNRPLCTATVNATNKMGFLNSQVN
15	1365	A	257	425	68	THAKFLNKKFNIPKLVILPKLVYIVKAIPTKM AIEFLLECDQNITVKLICENT*KNIAKNI*KRRV TFTPIET*HPVKQMIKWQ*LTAWLRNRGYKKI KQTPNSETAPSVCRNLVFDKCG
16	1366	A	263	104	481	FCIFRTTEEDRGGDDCVVSVWTKQRNNSCVK SKDVFSKPVNIFWALEESVLGVKARQPKPFFA AGNTFEMTCKVSSKNIKSPRYSVLIMAEKPV GDLSSPNETKYIISLDQDSVVKLENWTDASRV
17	1367	A	298	68	208	RKRTNNPIKLDKKFEHFKNEDI*ITSKHTKMW VSSLAMKEMLTKTTM
18	1368	A	300	904	1	LVVGITGTRHHARVIFIFLVETGFPHVGQAGL ELLTSGDPPALASQSAGITGMSHCARPKGHFG

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	[in	nuclcotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	Ì	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ļ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1	Ĭ	914	ng to first amino acid	acid residue of peptide	Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
	ł			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon.
	1			peptide .	sequence	/=possible nucleotide deletion. \=possible
		!		sequence		nucleotide insertion
-	 		 	sequence		IHLK*MFYTMSQKMP*PTINLILLLIIPGNLNIF
						KPNMGWLGPKTAFV*KDEVLSGIPFAKGRCR
	[1	i			WK*DY*C/LQEVTDPIMEKGKKKKRTASFFK
]				GQPHQSTNALLRRCVR*RYHLS\TVETAGLP*
		{	l I			KNTGHIPGQPFLFKLVFKC*NVICI**QYKW*O
	l	ł	1			NIGVKNKSFCPH*SSSPSL*FIGHHSRNF/CSFK
		ļ				TEPHSVVQAGGQWRNLSSLQAPPPGLMPLSR
		L				ISLMSSWDYRRPPQ
19	1369	A	302	3	445	NSPSRWAKIQMFEHTFCG*GCG/ER/NVHIHCS
		ĺ	i .			WICRLRPLLWRAVREYLSKLKNAELSFDPGV
1	Ì		{			SLLRIYAIDMPTSI*DEKEALLFAFLAFHE*HC
1						KSRIWAVIQ/CIHLWDWLRKL*CFHRMKFYA
20	1370		304	-,	1220	AV*NKPRHLLSHIWKDVQNILLK
20	13/0	A	304	1	1339	FFFCGKEVPLFEQNKHPGPRATTSPGA/HARA
1	 					LLSAGEFTAGVGLSP*AIHSFVWLCTFIQHGA
	•					GGPCHQPGGSPGPWMHTTQAGHLWEGAYPG
	(ĺ	!			GSSTWHQVPGQLGGSWGPRERSLLGSFIKCSP CPHPPGFRLWMSPNQKPPTENPGVMGRVWR
1	i			,		LMPGESPLIWEAEGKEDHLSPEGQGHSE/PVA
	,					PLHSSLGNTVKP*PKNQKPKONRSRHGO\GF
	•					MAGQGQSRPAAR*PPCPALTPASHSAGTWPP
						RICRTVPGGPCPSPSGFRSCRR*GFSA*TRSWP
						DAEPPSTPDTAPRCCTQSDTSSQGPQ*S*WRR
						CRALPGRLCSAPAAGLRRARPRLSESRRGNSP
	İ					PASPAAASARCPSWGPSCPARPPSRPAAGTEP
1						AAPSRCTAWLRGEREPGPRPPGRRPRSGRGP
						VSFAPEVLSLPAVRQTKSWRWRNEEEITRPW
						ALVRSRGG
21	1371	Α	326	799	1587	GSQVLPPRPSQDSATLPQDA*GPRAAPGQPVC
						E*GLQGAGVRRLRGEVLCQPQP*GAL*EQCLP
 						HLSFSPRQGAAPDTEPSAWGPAPTGATGPGLP
1			!			LRHVRLFSAGAPRGAATPCPPALLHGPAWPP
	1		}			ARPMFRGHPPVRPLGPWGKVAAGPRALCLA GVPAVQGECATKPSG*GL*PAHLRGPPGPEVL
			i l			QWHWQLSAGRDPVPAEDPPL*EGPLGPGGPA
	Į į]			AAQAEPGADPEPEDKDQAAESRPAGAMSLSA
]						QGSGPVGGQGLR
22	1372	Ā	327	146	652	PHLENPHPEHSFPGAPLT*STLSWSILSPREPSP
						GAPCYPGHPHLENPHLEHLLTWRTVTWSTLL
						PGAPCYPEHPHLEHPLTWSTPHLEHPSPGEPL
						SCRTPTRSILHRDHPLP*CLSTEESPI*GWGSLP
				İ		APPSTPLVLDVAPPGPQPASSCPGRDSCYSVP
						GTVVSP
23	1373	Α	348	397	2	CIVSSCQGTRKPCHLEDANKINKQSPTLEKIES
'						LQESL*VKQ*LIVAEKYVQILHPRKKYFQRPL
				ĺ		NNEKRKMKKRKEEKKKCRERMQRRSKWRR
				ĺ		EEKKE*RREE\EERKKEKEDRKERRKETSPRG
24	1374		362	170	250	SRRLLRD
Z4 .	13/4	Α	302	1/0	352	GRALDTAAGSPVQTAHGLPSDALAPLDDSMP
25	1375		384	373	128	WEGRTTAQWSLHRKRHLARTLLVSRVRGPQ YLITTILETGYLWKNRHSDQ*KRTENPERDQH
2.5	1313	^	204	دا د	120	VVDVVDECK SNEMVNDI CARVITIUM NUMBER DON
						KYPKVDFCKSNSMKNRLCNKWHWTNWIFTD
26	1376	A	397	383	165	KKINLNLKPHTKLTPNIKKN EVKNTNPFIFSGTNLTIWIRSI*RKSDEINQRTK
["	15,0	11	37,	203	100	*MEKYSISLDRRLNTVKMSFLPNLIYKFNTISI
		ľ	1		i	KIPANF
27	1377	A	406	103	380	KSKATGYMVNI*KLIV\FLYANDEQLEIEMNK
]				1		IVP\FNGSKNKIAFTNLTKYQNIQNRHAENYKI
j j		j	j	ſ	J	LVNKIEDLNKWRNVLLSWIGRRNIINTMT

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1.00	in in	nucleotide	location	F=Phenylalanine. G=Glycine. H=Histidine.
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	[09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
]	ļ	peptide		/=possible nucleotide deletion, \=possible
		_	}	sequence	}	nucleotide insertion
28	1378	A	408	14	427	TICTNKFNNLDEIK/FLERHKLSKLTQEEVENL
		ļ	ļ]	1	ITLKTSRETELVINK*VIPHKEKPGPDSFTGEF
						YQTFKEEL/II/ILHKLFQTIKYGRILPNSVYETSI
						TLKPKPEKDL\KENYRPLPLSNIDAK\LNKTLA
						NRI**HIR
29	1379	Α	434	395	128	IYSKMCMERQRLNN*ILKKNKVRGIAVPDVK
		i	-	!	1	VYYKPTVIK/TSWIL*KDSHIVEWNRLENLEID
						PN/IKRLILDKGAEATEWRKDSFFRQWQ
30	1380	A	455	2	228	FFFETESHSVTQAGVQWCNPGFKRFSCFGLSS
		}		ŀ	i	SWDYRYAPPRP\ANF*FLVETGFYYVAQAGL
						KLLSPGDLPALAS
31	1381	Α	462	393	2	QLMFDKGVKNIH\WGWTPPFTK*YWKNWISI
			1	ļ		CRRMNLNPYLSRYIKINSR\KDLTVRPEPIKLV
		j]	J	j	EENTGKTIQDTGLGK*FIAKTSKAQSTKTNK*
			1	ļ		KRQTRYIKLK\KKSTASKENNRVKRQPLE*EK
32	1382	-	474	125	471	IFAN
32	1362	A	4/4	125	4/1	VKPYEIAVFLVKPIEYK*HLLSDPAIPLSGI*LK
						EIKAYT/RRICTPMFAAPVSVIA/RN*KQSK/CQ KQ*YVHRMEYYTTIKRSEILICTTTWVDFRNT
	1		}	ļ	ł	ILRETDRIHKTTYDVISLI
33	1383	A	488	1825	72	KSACSFICSEEOPASPSPLKPGTYASEI\RPRDP
33	1303	Α	700	1023	~	HAAGPRRDSSEAETRRPRGA/DGSGTVVKGT
		1		}		PGSPAPPCSWGHGG\ETEGAG*CPAAPGTDLR
		ļ				APGGSAGS*\GLPSAGGSRGRKGWRAAGRQP
					ļ	STR*GRPGRHGGRGE*AGHPEPRQSALQSAG
		1				L/ASSPEPMGAALAEDGSGDSRGAGPRPQE*P
					1	PSVLSRS\GS*G*G*AASGTASSPRSHSSRLGPP
		1			İ	SAGFHGLRCGQPPFAAAPPGPWPGTGRPAGG
						AGSPPAAAGTAPPATRGAQSRRQNRTAGRNA
				l		SPQTAAGAGSPVQWALSRATG*TGETGSWC
		i	ł			AGGTHQATHLTAAWVCPPTWSVRPGGSGPA
						AGLGR*GRHPAQSPPLPVPRG*PAWPQEAPSP
			1			SPASSEVALSSGSCWPDQAPGPARGSPPAPLA
	'	[Ţ			PAWPAAGRGRQR*GRQSAHPPPRR*STAVSL
						SGTS*WRRSP*AGTRTQQC*SPWLVPACSSRP
		l				L*RGTRRPSTQQSPQTTGTPGRSAGPGHPRS*
						GGRSPAGTGHLGAQTVASPH*GHWPTALSCL
			1			WASASPPGPEAPPQTGACIGTNCRYRAASAR
						RSSVAPACA*GWQ*AGSPPAVLRGPP*RVRER
34	1384	A	497	422	2	GALTHRPRAPDE APGASVGRAQAAEG*RGGPTGRPPSALGVS/E
77	1304	1 ^	1 757	744		AGRAGRAGEGRPVPPAYPLCKSAQTSGPPKA
						RLS\PPLASCGGRGPPGGAACATCAPPAGPAR
		}	1			SSRCRRRSPPE*GPR*PSRPARPSPGSAASRRO
			[KLTPCRCOFRGLCA
35	1385	A	509	156	475	PTPYPGE*QAAFLLRGPGLRPPA/DPSLR/HRN
"	1505	l ' '	"	150	7/7	LTELVVAVTDENIVGLFAALLAERRVLLTAS
		Ì				KLSTLTSCDHAFCALLYPMRWEHVLIPTLPPH
		ĺ	[. 1		LLDYC*CPPLPRT
36	1386	A	512	3	1631	FFFSFVCHLYCVSPTPGPHGRLATWL/PGLLA
		1				FLGLAAGGQTLCPAGELPGHARAQASGAPGS
		[VLIAVPGRRRVHTCGPGPAAPSTRGECPPPAL
						GHTRPARPRPV\PFAPAVPQEPGGQGHGAA/P
)				PATGHSAPRGCPPARAAPTGSATPAPPPAACA
		i .	ı			
		l				Arnsawsyrraukuuurkyrararkkiirii
			,			AFHSAWSVPPAGRQQG*RVPAPAFRRTTPGT PGQHLLDRPGAPPAQGSGPAPAPPPRLAGPA
			į			PGQHLLDRPGAPPAQGSGPAPAPPPRLAGPA GPAAPPPGPPAASWHSSLSKSSSSL\GWSPPLP
						PGQHLLDRPGAPPAQGSGPAPAPPPRLAGPA

SEV ID	CEO ID	Mai	CEO	Dung!: at = 3	Dung!	I Amina cold comment (1) All 1 Constant
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
eotide	peptide seq-		in USSN	nucleotide location	location corresponding	F=Phenylalanine, G=Glycine, H=Histidine,
seq-	uence		09/496	correspondi	to last amino	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
uence	dence		914	ng to first	acid residue	
derice	ļ	1	914	amino acid	of peptide	Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
	}	ĺ	Î	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		1		peptide	Sequence	/=possible nucleotide deletion. \=possible
I	j	[1	Sequence	1	nucleotide insertion
	 		├──	Soquence	 	GRPSQSQ*PAGPPGMRGCCLRGW*PSSSGSD
						GPGPHPASTWLRAGKTGPSPPACGCA*LPPPS
	ĺ			ĺ	1	VSAAPQSPRTRCPRGCAAAAGLCVLAAAGAS
	1		ļ		Į	HGA\GLPGVRVHTORVHIH*GAG/GCOTPRPR
1	1	1		}	}	LRSLPVLGLPAPRCPVSAHPWHRRSGSSCHA
1	}	Ī	1	ł		ARLVPRHPAPGCP**TG*\PLITGFPEP*A*GLP
1			j	j		NHQAVGLEASGALQAGHRDELPTMVQLLDH
1	1	ŀ				SPDYPLKGRPHAP
37	1387	A	620	828	1	FRLPLAAGA/RGAAEPRVAVSMAPDPSAKIH
1 "	1.50	1 ^	020	020	1 *	WEASPEMQSKCHQKGKNNQTECFNHVRFLQ
İ		ĺ	1		{	RLNSTHLYACGTHAFOPLCAAIDAEAFTLPTS
1		l				FEEGKEKCPYDPARGFTGLIDGGLYTATRYE
i		İ	İ			FRSIPDIRRSRHPHSLRTEETPMHWLNG*EDE
						AQDDGG*GTISSFLLPWPADHPTPKSPGEPVH
1	ì	ŀ				SIPVCCQVRGQPQSGGKESPACLKSLSNCLTH
						\DAEFVFSVLVRESKASAVGDDDKVYYFFTE
1	ļ		[RATEKESGSFTQSRSSHRVARGIPPL
38	1388	A	739	1	427	FRAMVSSTLKLGISILNGGNAEVQ/QGNRGKG
1 50	.500	**	,3,	•	127	TSEEGKEG*EVPV*LPVSPPLPRPLQKMLDYL
1	}	ļ	}			KDKKEVGFFQSIQALMQTC\GEKVMADDEFT
!		ŀ			ł	QDLFRFLQLLCEGHNNDFQNYLRTOTGNTTT
1						INIIICTVDYLLRLQESI
39	1389	A	767	1	1030	TLDLTGPLLLGGVPNVPKDFRGRNRQFGGCM
"	.505	* *	'''	•	1030	RNLSVDGKNVDMAGFIANNGTREGCAARRN
						FCDGRRRQNGGTCVNRWNMYLCECPLRFGG
		·				KNCEQGEWPASSIPPVTAAWEALLLDVPGTT
		ŀ				VRGLHIQVRQPLVVYAAFTVDSHRPLQETVL
		ĺ				RRAPAPASGVPSPSGVGWDR*AGPAEPSPSTP
						ATVIISVPWYLGLMFRTR\KEDSVLMEATSGG
!						PTSFRLQVTGAPCHQGTC*VGARGRDPMLSG
						LRVTDGEWHHLLIELKNVKEDSEMKHLVTM
		· '				TLDYGMDQVSWHLHLLWG*TLPPAQGKTGA
j .]			j j			SEDKVSVRRGFRGCMQVRGGCGGRGEACPS
						QAAPRL
40	1390	A	801	69	399	IHKIIIHKEDLNKWKYILCSGMERLSTVMIPVV
						PQIIYKFNA*Q\VILKFTW*E*GAKITILRKNKL
						RGLVLVPLSTC*VKYLLDKVLPHIKTYYEAR
						VNKSVVLVQVTIM
41	1391	A	835	7	195	SMLKERKVFQFPSCLFFQYITWLGPPYHVLFD
}						SSVINFSIGAK*DILQSVMNCLYAKRIPCVT
42	1392	A	841	1	415	GSTHASGYDKTPDFILQVPVAVEGHIIHWIES
					-	KASFGDECSHHAYLHDQFWSYWNSLKHRTW
]						QGIGTVASNLSQL*TLNAPFPELLLFRSLARTG
.				i		FVLT*\RFGPGLVIYWYGFIQELDCNRERGILL
						KACFPTNIVTL
43	1393	A	845	358	92	PALSPAPVPQKKGSPLPLDPCLGPSSWLLSVG
				•		LGWPRL*PRRGPGDPGSLPATPPLLTPPHTLLP
]			QRPMLPPSHAGLARPPPPEPISVP
44	1394	A	853	452	1	LPOYCFFPRLSPKSKLVKHSAL**PSALKPPTK
.					-	SPRCIPRTSLYFTICC/PPALQL/SPIEDPPAIYRS
		ľ	1	ļ	ł	PPTHMLRSASQPLNQAPTLVKGHPPSRFLQG
			l	ĺ		QVSCPPQPTLPREKPLPLHLRPPPRPAQPPLPR
			1	i		PLTFSTRRNVDPEIPERFR
45	1395	A	894	379	162	GVYPPTVFDNYSVQTSVDGQIVSLNTWDTAG
'-	.020				- 52	QEEYD/RLRTLS*PQTSIFVICFSIGNLEFPIYGT
j			j	1		WLSMSMGK
46	1396	$\overline{\mathbf{A}}$	900	1	366	TTKKTLISNNVSSRSLPILPELKAFSLAFNDPL
.		· 1		-		EIQKYMRT/DQ*CVTHDISLYIVTKLALIFLIPR
		1	i			VFLFHQLNIT**CLHFFTMTTFIAIPFSFLFLGR
		1				TELLIQUITE CELLITIVITE ITALIFOI EFEOR

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion D/KSLAMLPRLVSNSWPQVILPP QLQNLASRGCL*SQLLRRLRRENRLNPGGGG CSEIAP\CTPAWVTQRDFFRKKK HFTPDRIAIVKNTRDSHCWRGC*EEGAPARC PRKRESWWGERLP/PRGFPPAAEDAPAPGWK
						GRKHASRTARAHVFHPIRQSIRSPVRGRPGDP RAAHTRSAGTRLQCKASRGG*GKGPAPTR*E GGPGSAPAPLPASSGCSLFPDSSPWTPPPPAPG AAAAQP**TPRCPAALRAGAHIGRVGRPY
50	1400	A	973	45	421	EKCIQALDVFVFCYIDHSSHCLMSCD*E/DQA LNFMPLEMEPKMSKLAFGCQRSSTSDDDSGC ALEEYAWVPPGLRPEQIQLYFACLPEEKVPY VNSPGEKHRIKQLLYQLPPHDNEVRYCQSLSE E
51	1401	A	992	2095	194	IRIRHEAARSCLGCAAGHVPAPGLRLLPTVRG PPGRRGPAAPGCVCY*SGESTFVSHVPQRMA WPGSAPPRGFHPLQSQTSPSDTVSSPQLSKEE DGPGWEHPLSSSL*SLGQAGGNH*QPEELAG WEPRGPPSLAPSSPT/TMWTALVLIWIFSLSLS ESHAASNDPRNFVPNKMWKGLVKRNASVET VDNKTSEDVTMAAASPVTLTKGTSAAHLNS MEVTTEDTSRTDVSEPATSGVAADGVTSIAPT AVASSTTAASITTAASSMTVASSAPTTAASST TVASIAPITAASSMTAASSTPMTLALPAPTST STGRTPSTTATGHPSLSTALAQVPKSSALPRT ATLATLATRAQTVATTANTSSPMSTRPSPSKH MPSDTAASPVPPMRPQAGGPISQVSVDQPVV NTTNKSTPMPSNTTPEPAPTPTVVTTTKAQAR EPTASPVPVPHTSPIPEMEAMSPTTQPSPMPYT QRAAGPGTSQAPEQVETEATPGTDSTGPTPRS SGGTKMPATDSCQPSTQGYMV/DHH*APHP GRGRQNSPSGGAVTRGDPFHHSLGFVCPAGL *ELQEEGLHPGGLLNQRDVCGLRNVRGAGA WREAWPLPRPFLLPLRPNQVLPNSFGAIEEIC QMLKHI
52	1402	A	994	1	462	ESGEFLVSFTLKKPTNVFHHINGMKFFNK/LIF *SHTDIAFYKIQHPFMLKALTKWA*EGT*PDR RYLH*SLRLNGEQLKTFPLRSGMR*G/CAILPL VLNAMLSIVPAVVPAGKTRHEKEITCPLIGQE EK*FS*FVGDMNTCVENKKESKKLLE
53	1403	A .	1011	1	630	PEVIQQSAYDSKADIWSLGITAIELAKGEPPNS DMHPMRVLFLIPKNNPPTHCWRRLLESFKEV *LMLA*TKDPSIRPTAKELLKHKFIVKNSKKT SYLTELIDRFKRWKAEGHSDDESDSEGSDSES TSRENNTHPEWSFTTVRKKPDPKKVQNGAEQ DLVQTLSCLSMIITPAFAELKQQDENNASRNQ AIEELEKSIAVAEAAGPG
54	1404	A	1016	1	222	ISIDA*KAFDKIQH/CFMITTLKKLGIDGKYLN TIKAIDDRHTVSTILNVEKLKAFL*RSGTRQRF PISGSGARI
55	1405	A	1033	3	366	HASVDGDEGSDDVYYYYTPAILRELQALNTA EAAEHRPEEDRMLSEDPWRPAHMIKGYMPL HNIPHTEVIDVTGLNQSHLYQHLNKGTPMKT QKRAA\LYTWHVLEQLEILRQINQQSHGPG
56	1406	A	1044	5	429	SVLTLQTRSPSKPLS\RKLMDWEVVSRNSISE DRLETQSRASRSPPVTPNQSQETPVDGKPLAL PPNQSQKNIRYHIHYLHLQYYLDRHISATLPIP SSSGIPTPIAVITDALTDLVELILGQPCSEESGR APGTLFLLAL

CECID	CEO ID	1 3/04	Tero	Deading	David and	Lanin and annual (Analysis Co. Contribution
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleonde	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-)	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
ì				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	{	l	ł	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		Ì		peptide	Sequence	/=possible nucleotide deletion, \=possible
i	l	1	ł		1	
		<u> </u>	 	sequence		nucleotide insertion
57	1407	A	1050	11	430	GAYAFETNGFPIMLVLTTDKIEGDVGIAGLYD
1	}		l			MH\ISLPMAFLLRTLVRCTSYIIPVTHVLSTPV
1			f			TCLRRREKDGVIVDVLSDTASNHNGFPVEEH
	!		Į.			ADDTHPARLOGPTLRSQPMGPLKHKAFEERA
		1				NLGLVQRRLRLED
58	1408	A	1058	258	419	LKHRDTPVVGANNRALSCTPLTSLTLCALCPL
70	1400	Ι ^ .	1038	236	419	
						PCLGCPTXATCRLYQTTVAVVF
59	1409	A	1064	3	425 ·	KAFSFTTSLIGHQRMHTGERPYKCKECGKTF
			1			KGSSSLNNHQRIHTGEKPYKCNECGRAFSQC
1	1	l	!		· ·	SSLIQHHRIHTGEKPYECTQCGKAFTSISRLSR
1	ł		ł		ļ	HHRIHTGEKPFHCNECGKVFSYHSALIIHQRIH
Į.	!	ł			Ì	
		ļ. <u>. </u>				TGEKPYACKDVGK
60	1410	Ā	1065	204	419	GGPPGPFLAHTHAGLQAPGPLLAPAGDEGDL
	1	(1		1	LLLAVQQSCLADHLLTASWGGK/DPIPTKALG
1					l	EGQEGLPLTV
61	1411	A	1079	3	383	RHSRAHLCQPFHLVMRDLLQLGQDIPQGCHY
1	l	١	1,	ا آ		LEENHLIHRDIAARNCLLSCAAPTRAATIGDF
1	ł	Ì	1		ļ	
1	l		1			GMARYIYRTRYYQLGDRAL/LPRKWMPPEAL
	[1	i			LEGIFTYNTDSWTFGVLLWEIFSLGYMPYPGR
1	l					TN
62	1412	A	1080	1	859	VVEFLWSRRPSGSSDPRPRRPASKCOMMEER
			1	_		ANLMHMMKLSIKVLLQSALSLGRSLDADHA
	İ	l	ļ		İ	PLOOFFVVMEHCLKHGLKVKKSFIGONKSFF
1 .	J]	ļ		J	
1 .			ļ		•	GPLELVEKLCPEASDIATSVRNLPELKTAVGR
	ļ		ł			GRAWLYLALMQKKLADYLKVLIDNKHLLSE
	ļ	ł			1	FYEPEALMMEEEGMVIVGLLVGLNVLDANL\
1 .	į		ļ		ļ	CLKGEDLDSQVGVIDFSLYLKDVQDLDGGKE
1		1	1			HERITOVLDOKNYVEELNRHLSCTVGDLOTK
1	}	1	1		ļ	IDGLEKTNSKLQERVSAATDRICSLQEEQQQL
		-,				REQNELIR
63	1413	A -	1083	2	615	SSFAKHKRIHTGEKPFICLECGKAFTSSTTLTK
i i		l	{			HRRIHTGEKPYTCEECGKAFRQSAILYVHRRI
		ŀ				HTGEKPYTCGECGKTFRQSANLYAHKKIHTG
			1		ļ	EKPYTCGDCGKTFRQSANLYAHKKIHTG\EKP
		1	1			YKCKECGKAFKSYYSILKHKRTHTRGMSYEG
4	[ĺ	ĺ			
		1				DEC/QRSLN/RSSILSNHKIIHNEEK/PLKCEKCE
<u> </u>			<u> </u>			KAFNHTSICCRHKKN
64	1414	A	1084	946	1 .	KKQDLSSSLTDDSKNAQAPLALTESHLATLA
		1				SSSQSPEAIKQLLDSGLPSLLVRSLASFCFSHIS
]		ļ				SSESIAQSIDISQDKLRRHHVPQQCNKMPITAD
1]			LVAPILRFLTEVGNSHIMKDWLGGSEVNPLW
j i		J	}		1	TALLFLLCHSGSTSGS\HNLG\AQQDQCKISFS
1			[
1			1			FFSWLTTGLTTQQRTAIE\NATVAFF\LQCI\SC
]				HPNNQKLMAQVLCELFQTSPQRGNLPTSGNI
1			1 :			S\GFIR\RLFLQLMLEDEKVTMFLQSPCPLYKG
1] [RINATSHVIQHP\MYGAGHKFRTLHLPVSTTL
1						SDVLDRVSDTPSITAKLISKQKDDKKKK
65	1415	A	1087	103	324	PRAFEFVHTEMIVG/RVQNIHLFTLQVLEDRA
1 00	1713	Λ.	1007	103	J24	
, I]			LFTMSVGSSLWSTYLIHVMALP/DRELLKPNA
			L			SVALHKLSNALV
66	1416	Α	1095	3	493	HETCSVTHIVSFSLPFLNPSHPASTPGHTENEQ
1 1						PSLVWFDRGKFYLTFEGSSRGPSPLTMGAQD
1 1						TLPVAAAFTETVNAYFKGADPSKCIVKITGE
						l I
1						MVLSFPAGITRHFANNPSPAALTFRVINFSRLE
						HVLPNPQLLCCDNTQNDANTK\EFWVNMPNL
L			[MTHLK
67	1417	Α	1098	57	356	LKLTSLGFIIGVSVVGNLLISILLVKDKTLHRA
j l				ļ		PYYFLLDLCCSDILRSAICFPFVFNSVKNGST
j				ĺ		WTYGTLTCKVIAFLGVLSCFHTAFMLFCISVT
	LJ					

SEO ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine.
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine.
eotide	seq-	ł	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	Ì		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1		ĺ	ĺ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1		İ	ļ	peptide		/=possible nucleotide deletion, \=possible
1		1		sequence	ì	nucleotide insertion
 		 	 	1	 	RYL
68	1418	A	1106	1	1326	MGKISATGINMGTKCSWALVWHLESYDPKH
""	1	**		· .	1520	YEREGMQDWKTASGQSEEATQQSSQKPOPH
ĺ.		ì	i	ĺ	Ì	YTTYQSSSFLKYSSESHLLAWRENSSEGSFQF
1		l	ļ	i	{	PGRSRARPPRTRQQRRGAAAGPGRGAVRLG
1			1			HPQSAAQPQLRAAARIPESPAAFPAQPRPGSA
		ļ		١.		RNSDASGPASLSRTLGRASSPRPPQAPDVTAP
1		}	1			SPAALAPRAARGGSRAAALAGAEAEEPLRTL
						APRPTRAAAPPPPPPPPPPLPPGAPPPPVRCVSR
		Ì				RARAPPWR/PAATGPPP\RPVAPSRKLGSARAP
1						1
1 .						APALQIRKGTSSGLPGRGGGSGPGNNLSSVA GNWRGSSFAVERPGMAKYQGEVQSLKLDDD
'						1
						SVIEGVSDQVLVAVVVSFALIATLVYALFRNV
1		1				HQNIHPENQELVRVLREQLQTEQDAPAATRQ
1						QFYTDMYCPICLHQASFPVETNCGHLFCGSLT
69	1419	A	1107	2	466	PNSIW
09	1419	A	1107	²	400	FDTARLHEFGTSITQIFAVDNREDLQKWMEA
						FWQHFFDLSQWKHCCEELMKIEIMSPRKPPLF
1 1						LTKEATSVYHDMSIDSPMKLESLTDIIQKKIEE
						TNGQFLIGQREESLP/SS/CGPHSLMVTIKWSS
70	1400					RKRY/SYPASEPLHDEKGKKRQAPLPPSDK
70	1420	Α	1111	698	23	ALRRLHYVRATKV\FLSFRRPFWREEHIEGGH
						SNTDRPSRMIFYPPPREGALLLASYTWSDAAA
						AFAGLSREEALRLALDDVAALHGPVVRQLW
						DGTGVVKRWAEDQHSQGGFVVQPPALWQT
1				!		EKDDWTVPYGRIYFAGEHTAYPHGWVETAV
						KSALRAAIKINSRKGPASDTASPEGHASDMEG
						QGHVHGVASSPSHDLAKEEGSHPPVQGQLSL
71	1401		1110			QNTTHTRTSH
/¹ .	1421	Α	1119	2	385	QKQTLQNGYLDSSMDILYLGSLPPELQVSSDE
1 1						PPGPPEQAGLSQFIILEPETQNPETTEEIQSS\LQ
						QEAAAQLPQLPEVVELSSTKA\EAPALPSQSL
	1400	_,				EGVHSSTEQKAPAQQLPAFEEILAPLLIHHE
72	1422	Α	1127	1	906	HAQYVGPYRLEKTLGKGQTGLVKLGVHCIT
ĺ				ĺ		GQKVAIKIVNREKLSESVLMKVEREIAIL\RLI
! !	·					EHPHVLKLHGVYENKKYFPPDELTSGPSMLA
! !		ļ				QVSPHGKLSARRSWDLLSGFPRYLVLEHVSG
ļ <u> </u>	j	- 1	ļ			GELFDYLVKKGRLTPKEARKFFRQIVSALDFC
]]						HSYSICHRDLKPENLLLDEKNNIRIADFGMAS
			{	. 1	i	LQVGDSLLETSCGSPHYACPEVIKGEKYDGR
1			1			RADMWSCGVILFALLVGALPFDDDNLRQLLE
1 1	ļ	}		.	Ì	KVKRGVFHMPHFIPPDCQSLLRGMIEVEPEKR
172	1400	-,l				LSLEQIQKHPWYLGGNFIS
73	1423	Α	1128	1	802	LRNALDVLHREVPRVLVNLVDFLNPTIMRQV
			,	}		FLGNPDKCPVQQA/MLEPLGSKTETLDLRAE
	1	l	}		i	MPITCPTQNEPFLRTPRNSNYTYPIKPAIENWG
[į	ļ	ĺ	SDFLCTEWKASNSVPTSVHQLRPADIKVVAA
	ļ	ļ]	J	ļ	LGDSLTTAVGARPNNSSDLPTSWRGLSWSIG
]	j				GDGNLETHTTLPNILKKFNPYLLGFSTSTWEG
		i	1		[TAGLNVAAEGARARDMPAQAWDLVERMKN
		ļ				SPDINLEKDWKLVTLFIGGNDLCHYCENPEA
<u> </u>						HLATEYVQHIQQALDILSE
74	1424	A	1139	60	480	FREPCLLVPGDHQPLREASWLA/LPPIGLWGT
						DSPLCCVEVAIPCNKGAHSVGLKGWLLAQG
1		ļ	!			VLGMRDTIPQEHPWESTPDLCFCRDPEEIEVE
	1	1	i			EQPAADAAVAKGEF/QGEQIAPVPA\IIAAHPE
	}	ł	1	i	ł	AADPAPVHTTAHPKGA
75	1425	A	1147	2	413	PFPHQHPQEP\KGSCWPQSALRGQCPGPVLGV
		[i	Ì		TTTSDLCSLQVPVSSHRNPLLDLAAYDQEGR
						<u> </u>

SEO ID	SEQ ID	Met	LOTO	D. J. A. J	100 - 100 - 100 - 1	[A ::
		1	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	,	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1	}	1	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
ł	J	ļ	}	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			İ	peptide	Soquence	/=possible nucleotide deletion, \=possible
ì	l	1	ł			, . , , , , , , , , , , , , , , , , , ,
ļ		<u> </u>	<u> </u>	sequence	ļ	nucleotide insertion
	ł	į.				RFDNFSSLSIQWESTRPVLASIEPELPMQLVSQ
			1			DDESGQKKLHGLQAILVHEASGTTAITATAT
	į	ĺ	i	ĺ	ì	GYQESHLSSAR
76	1426	A	1155	38	410	PIISAPAQDDPILLSFIHCLHANLLCVWRRDVK
1		l	Ì		!	PDCKEIWIFWWGDEPNLV\VQYIMNCMLWK
	!	}				KDSGKMAFPMNVGRC/FFKEIHNLLERCLMD
	ſ	l	[[[
					į	KNFVLIGKWFVRPYYKDEKPVNKSEHLSCAF
						T
77	1427	A	1162	526	350	RFPQGLEDVSTYPVLIEELLSRGWSEEELQGV
	1	!		ł	ŀ	LRGNLLRVFRQVEKVQEENKWQSPLED
78	1428	A	1171	I	1293	MAESASPPSSSAAAPAAEPGVTTEQPGPRSPP
1	1		1	l	1	SSPPGLEEPLDGADPHVPHPDLAPIAFFCLROT
Į '	l	1	1			TSPRNWCIKMVCNPWFECVSMLVILLNCVTL
		1		}		
	ĺ	1	(GMYQPCDDMDCLSDRCKILQVFDDFIFIFA
	l	1				MEMVLKMVALGIFGKKCYLGDTWNRLDFFI
j		ļ	j	l	ļ	VMAGMVEYSLDLQNINLSAIRTVRVLRPLKA
		l			1	INRVPSMRILVNLLLDTLPMLGNVLLLCFFVF
		1				FIFGIIGVQLWAGLLRNRCFLEENFTIQGDVAL
1		}	!]	PP\YYOPEEDDEMPFICSLSGDNGIMGCHEIPP
.}	i	İ				LKEQGRECCLSKDDVYDFGAERQDLNASGL
1			[, -
			ļ			CVNWNRYYNVCRTGSANPHKGAINFDNIGY
) .	j	ļ)			AWIVIFQVITLEGWVEIMYYVMDAHSFYNFI
						YFILLIIVSVREPGLLGGSFSTAQSPKCQGDSFP
1						GVAAESLLLRGWVLWLPGGG
79	1429	A	1175	1	405	PNDFFKDMFPDLPGGPLGPIKAENDYGAYLN
1 1			1			FLSATHLGGLFPPWPLVEERKLKPKASQQCPI
]			CHKVIMGAGKLPRHMRTHTGEKPYMCTICE
						VRFTRQDKLKIHMRKHTGERPYLCIHCNAKF
1 1			i 1			
-						VHNYDLKNHMR
80	1430	A	1182	25	198	EMNELSQQLSQQGGRGASQCPSPPAPTLPNPT
						PLCQLQLQRVNTGLPTPPCHPGAGAA
81	1431	Α	1186	254	583	KTVLDVGAGTGILSIFCAQAGARRVYAVEAS
1 1			i			AIWOOAREVVRFNGLEDRVHVLPGPVETVEL
1						PEQVDAIVSEWMGYGLLHESMLSSVLHARTK
						VVKDGGFFLPXSSELFM
82	1432	A	1187	2	716	DFVDAARNLPLESTKSPAEPSKSVPSLE\DPRA
\ \frac{\sigma_{\sigma}}{\chi}	1736	^	110/		/10	
1						SSQGLPSQGPVQNQGRRGEQRPKKF/TVIQHT
]					J i	SSFEKSDSLEQPSGLEGEDKPLAQFPSPPPAPH
1						GRSAHSLQPKLVRQPNIQVPEILVTEEPDRPD
]						TEPEPPPKEPEKTEEFQWPQGSQTLAQFPVEK
1 1						LPPKKKRLGLAKMAQSSGESSFESSVPLFRSP
j J						SQESNVSLSGSSRSALFERDDHGKAEAPSPSF
] [,	DMGPKPLGTHMLTV
83	1433	Ā	1188	517	904	
103	1433	А	1199	517	804	ESPGLSKVLRTGAFAYPFLFDNLPLFYRLGLC
						WGRGHGCGQEALSTSHGYHLFCALLTGFLFA
						SHLPERLAPGRFDYIGHSHQLFHICAVLGTHF
<u>L_</u>						Q
84	1434	A	1192	45	476	LGDVGFWVERTPVHEAAORGESLOLOOLIES
			I	_	· · · ·	GACVNQVTVDSITPLHAASLQGQARCVQLLL
						AAGAQVDARNIDGSTPLCECLRLGQHRVCEA
1 1			· •	ł		
1				į	1	LAVLRGQGQPSPVHSVPPARGLHXREFRMC*
						GFLFDVGXNLEAHEFHFGEP
85	1435	A	1194	69	410	KRSEEASAPPFPLGGTGAAPTRASLPEQILLPR
				l		SCLEARKSQPDEKLLSALHNSRTWN*EPRRSQ
1 1				}		HRLVSPEVHPGRRGSSPGVAECKLTSAYFRT
, ,						GRSPCPSLPGTTRTNSLL
96	1426		1215	3	405	
86	1436	Α	1415	ا د	405	LPSHTCGNPGRLPNGIQQGSTFNLGDKVRYSC
1 1						NLGFFLEGHAVLTCHAGSENSATWDFPLPSC
L l						RADDACGGTLRG/AEWHHLQPPLPLG/ATKN

No. of N	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Docation Docation	NO: of		hod	1 .		L .	1
Sequence 19496 corresponding to get the service of peptide residue of peptide sequence comparison of peptide sequence comp	1	,	1	1	1		
Part Part				1			
mino acid residue of peptide sequence physics of sequence physics of sequence physics of sequence sequ		uence					
Principle Prin	uence		[914			
peptide sequence		1					
	ĺ	1	ł			sequence	
NADCTWTILAELGDTTALVFIDFQLEDGYDFL	'	ľ	1				
87			ļ		sequence		
1437 A 1216 226 964 GTARFGPMYGFGANRRAGRI PSLVLGVILG	1	į	Ì				
	07	1427	-	1016	224	1007	
RTEVARGRLEKRNSDLFA VVGHAQETDRPS	°′	1437	A	1210	226	964	
	-{		ĺ			1	
SYQMADHILKEQLAELRQEFLRQEDLQA SYRNNYTLYKRLEVESFQCQQOMKELRAQH EENIKKLADOFLEEQKORTOKIQSNDCKELDA EENIKKLADOFLEEQKORTOKIQSNDCKELDA SERIKKLADOFLEEQKORTOKIQSNDCKELD NNQVVPKNIPK VAENNADKNEEPSSNHIPHO NNQVVPKNIPK VAENNADKNEEPSSNHIPHO RYKSAGAWIHIPYSDFPYWDLIMLLLMVGN LIVLPYQITFFKEENSPWIVPNYLDEDTFILD LVLNRRIGIVVEEGAEILLAPRAIRTRYLRTN FLVDLISSIPVDYIEVVELEPTLOABEVYKTAW ALRIVERTKILSLURL ALIVERYKILSLURL CELVEAVDGKYGMLTRSNAAPGRHLAMLET LVVVARRVDADMILINPOTLSLABNITVV APMLDSRAAYSNFWCGMTSQGYYKRTPAV APMLDSRAAYSNFWCGMTSQGYYKRTPAV PIRKRDRGCFA VPMYHSTFLIDLRAKASRNI, AFPPHPDYTWSDDIUVAFSCKQAEVOMY VCNKEEYGFI.PVFLARHSTLQDESSHMIVO LEVMYPSSPSSAQSMAVVSADHIGLVISVI, NRTSFFYLRNIVVADLIMILTFFRIVHDAGF GWDRFFILCRYTSVLFYANMDTSIVVLGLIT/ YDRYWKVVRHLWDSWMTGISFTRYYLLG GWDRFFILCRYTSVLFYANMDTSIVVLGLIT/ YDRYWKVVRHLWDSWMTGISFTRYYLLG GRANDWFGLLAKAGGHGGISWA A 1245 3 1937 LGSSDVRAPQRSELGAESPSRNVASQAYNIL SALTPHLTRSRVLNEEPLTLAGGSRAPANLSD VVQLIFLVDSNFPFGYISNYTVSTKVASMAR GRISSANSVVQPQAFVGAVVTLDSSNFAAV ALIVENTALGRAFITVENEPLTLAGFSRAPANLSD VVQLIFLVDSNFFFGYISNYTVSTKVASMAR GRISSANSVVQPQAFVGAVVTLDSSNFAAV LLQLNYTLLCGRYLSEPEPLYLAYVLHSEPP PNEINCSASRRIRPESLQGADHRPYTFFISPGT RDPYGSYRLALSSHFRWSALEVSAVLTSLC QYFSEEDVWRTEGLLPLEETSPRQAVCLTR LTARGTSLFYPPSRFWFFWFFETSHOTYNIYMIN LYGVDSRSGIFHLDGDRAFIRINDJIPQLATSLC QYFSEEDVWRTEGLLPLEETSPRQAVCLTR HLAAFGTSLFYPPSRCJFTRQRACLYSLC QYFSEEDVWRTEGLLPLEETSPRQAVCLTR HLAAFGTSLFYPPSRCJFTRQRACLYSLC QYFSEEDVWRTEGLLPLEETSPRQAVCLTR HLAAFGTSLFYPPSRCJFTRQRACLYSLC QYFSEEDVWRTEGLLPLEETSPRAAVCLTR HLAAFGTSLFYPPSRCJFTRQRACLYSLC QYFSEEDVWRTEGLLPLEETSPRAAVCLTR HLAAFGTSLFYPPSRCJFTRQRACLYSLC QYFSEEDVWRTEGLLPLETSPRAARLSDIFQLATSLC QYFSEEDVWRTEGLPLETSPRCSFTROKSCITAHVGMI LYGVDSRSGIFTRABLITEQLTATIL CANADWYGAVGDSAYSTGRVSRLINLSUDDIP HANDRIT SARTINDARSHMITHTYTCT SARTINDARSHMITHTYTCT GANAWWGAVGDSAYSTGRVSRLINLSON TYPPPPFGGFSFARALRKESTITSTGLEKEEL PGRAASCSTAGSGSRGLPPSSPMVSSAAHWIN SALPPPNYTAALLKERRETVAIDRLORGKNINLOND	1		1			1	\
				j			
EENIKKLADOPLEEOKOĞTÖKIĞSINDÖKELDI NNQVYPKINIPK VAENVADKIEPSSINHIPRIG	1		ļ	1]	
NNQVYPNIPK VAENVAENKAVEEPSSNHIPHO	1					1	
1438					1		
RYKSAGAWIHPYSDFRFYWDLIMLLLMVÖN LIVLPVGITFKEENSPPWIVFNVLSTFFILLD LVLNFRTGIVVEGGAEILAPRAIRTRYLRTW FLVDLISSIPVDYIFLVVELEPRLDAEVYKTAR ALRIVEFIKISLERL 89 1439 A 1223 I 743 MGFDEVFMINLRRRQDRRERMLRALQAGEIE CRIVEAVDGK VGMLTRSNAAPGRHLAMLET LVVVAPRFVDADNILLNPDTLSLLIAENKTVV APMLDSRAAVSNFWCGMTSQGYVKRTPAYI PIRKBRRGCRAVFMVHSTFLDLRKAASRNL VAFYPPHPDYTWSTDDIVFAFSCKQAEVQMY VCNKEEVGFI.PVPLRAHSTLQDEAESFMHVQ LEVMYPSSPSSAQSMAVVSADHIGLVISVI. 90 1440 A 1227 2 349 NKTSFFFYLKINVVADLIMTLTFPFRIVEDAGF GPWDFKFILCRYTSVLFYANMDTSIVVLGLIT/ YDRYWKVVRHL/WDSWMTGI/SFTRVYLLG GARLWYGKLILAKGGHGISWL 91 1441 A 1245 3 1937 LGSSDVRAPQRSELGAESFSRMVASQAYNLT SALTPILTRSRVI.NEPLAGGSGTAPANLSD VVQLIFLVDSNPFPFGYISNTVTSKTVASMAF GTOAGAGUPERLASBRATIVKVPNISDWAAR GHRSSANSVVUPQAGFVGAVVTLDSSNPAAV LHLQLNYTLLDGRYLSEEPEPYLAVYLHSEPR PNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDFVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVWRTEGLLPLETSPRQAVCLTR HLTAFGTSLFVPPSHRFVFPEFTANDVITVML TCAVCLVTWYMMALHKILDQLDASRGRAP PCGQRGFKYEILWKTGWGRGGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKRVWHDNKGLSPAWELGHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRYPTSNAALLERFRLLVABELQRGF FPKHIWLSIWDRPPRSCTTRIQRATCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPILSV DTVAVGLVSSVVVYPYLAILFLFRRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCKKLVATMP LFANADPINFYTAMLSKLRFEVFQPGDYIREG AVOKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRASAFPCPPLSSAAAR GKGVHQRLLGKGQHT 70FGEILLTKGRRTASVRASAFPCPPLSSAAAR GKGVHQRLLGKGQHT 70FGEILLTKGRRTASVRASFELKVATMP LFANADPINFYTAMLSKLRFEVFQPGDYIREG AVOKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLTKGRRTASVRASAHPPN KABIPERRKDSTSTINNLIPSPMMTRNTYVCT 83 1443 A 1249 180 901 TVPPPFGGFSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPRSSFMVSSAHNPN KABIPERRKDSTSTINNLIPSPMMTRNTYVCT	00	1420		1010		 	
B9	00	1438	A	1218	1	1 234	PEFGITISCGYLMATDVSRRPSVHKAVEIEQE
LVLNFRTGIVVEEGAEILLAPRAIRTRYLRTW FLVDLISSIPDVIPILVVELEPRLDAEVYKTAR ALRIVRFTKILSLIRL SP	1		ĺ				
FLVDLISSIPVDYIFILVVELEPRLDAEVYKTAR ALRIVRFTKILSLIRL	1					1	(
ALRIVAFTIKI, SLIRL	1						
1439 A 1223 1 743 MGFDEVFMINLRRRQDREERMIRALOAQEIE CRLVEAVDGKVGMLTRSNAAPGRHLAMLET LVVVAPREVDADNILINPDTLSLLIAENKTVV APMILDSRAAYSNFWCGMTSQGYYKRTPAYI PIRKRDRRGGCAPVPWISTEDILDRKAASRNI. AFYPPHPDYTWSFDDIIVFAFSCKQAEVQMY VCNKEEYGFLPVPLRAHSTLQDEAESFMHVOY CLEWMYPSSPSAQSMAVSADHIGLVISYL LEWMYPSSPSAQSMAVSADHIGLVISYL DVARPQRSELGAESPSRMVASQAYNLT SALTPILITESRV LNEGELTLAGPSSRAPANLSD VVQLIFLVDSNPPFGYISNYTVSTKVASMAF QTQAGAQIPIERLASSERAITVKVPNNSDWAAR GHRSSANSVVQPQAFVGAVVTLDSNPAAV LHQLNYTLLDGRYDSSPAPAVLHSPR PNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDPVGSYTRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVWRTGCILPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRPVSPEPTLADVNIVML TCAVCLVYTMVMALILLDQLDASGGAIP HSIGSMWIRWWHDNKGLSPAWFLQHIIVRD LYGYDSSVVYPYVTLAILDI, FRINSSKOF LYGYDSSVVYPYVTLAILFI, FRINSSKOF GRGVVHQRLJGKGVHTVALIFI, FRINSSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVVHQRLJGKGWHT LFANADPNFVTAMLSKLRFEVFQPGDYIREG AVGKMYFIQHGVAQVITKSSKEMKLITDGS YFGEICLLTKGRRTASVRADTYCRLVSLVD NFNEVLEEYPMMRRAFETVALIDGS YFGEICLLTKGRRTASVRADTYCRLVSLSVD NFNEVLEEYPMMRRAFETVALIDGS YFGEICLLTKGRTASVRADTYCRLVSLSVD NFNEVLEEYPMMRRAFETVALIDGS NFNEVLEEYPMMRRAFETVALIDGS NFNEVLEEYPMMRRAFETVALIDGS NFNEVLEEYPMMRRAFETVALIDGS NFNEVLEEYPMMRRAFETVALIDGS NFNEVLEEYPMMRRAFETVALIDGS NFNEVLEEYPMMRRAFETVALIDGS NFNEVLEEYPMMRRAFETVALIDGS NFNEVLEEYPMMRRAFETVALIDGS NFNEVLEEYPMMRRAFETVALIDGS NFNEVLEEYPMMRRAFETVALIDGS NFNEVLEEYPMMRRAFETVALIDGS NFNEVLEEYPMMRRAFETVALIDGS NFNEVLEEYPMMRRAFETVALIDGS NFNEVLEEYPMMRRAFETVALIDGS NFNEVLEEYPMMRRAFETVALIDGS NFNEVLEEYPMMRAFETV]	
CRLVEAVDGKVGMLTRSNAAPGRHLAMLET LVVVAPRFVDADNILINPDTLSLIAENKTVV APMLDSRAAYSINFWCGMTSQGYYKRTPAYI PIRKRDRRGCFA VPMVHSTEILDLRKAASRNIL AFYPPHPDYTWSFDIDLRFAFSCKQAEVOMY VCNKEEYGFLPVPLRAHSTLQDEAESFMHVQ LEVMVPSSPSSAQSMAVVSADHIGLISTYL 90 1440 A 1227 2 349 NKTSFIFTYLKNIVVADLIMTLTPFFRIVHDAGF GPWDFKFILCRYTSVLFYANMDTSIVVLGILTY YDRYJWKVVRHL-WDSWMTGUSFTR VYLLG LGARLVWFGKLILAKGGHQGISWL 91 1441 A 1245 3 1937 LGSSDVRAPQRSELGAESPSRWASQAYNLI GARLVWFGKLILAKGGHGGISWL UGSSDVRAPQRSELGAESPSRWASQAYNLS VVQLIFLVDSNPPFGYISNYTVSTKVASMAF OTQAGAQPIERLASERAITVKVPNNSDWAAR GHRSSANSVVQPQAFVGAVVTLDSSNPAAV LHLQLNYTLLDGRYLSEEPEPYLAVYLHSEPR PNEHNCSASRRIPPESLQGADHRPYTFIPSPGT RDPVGSVRLALSSHFRWSALEVSVGLYTSLC QYFSEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRPVPPEPTADVNTVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGRIRHLDGDRAFHRNSIDJFQIATP HSLGSMWKIRVWHDNALBFARGLPAWFLQHIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASRVPTFSAALLRERRLLVAELQRGF FDKHWLSIWDRPPRSCTFIRQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGYHQRLLGRGQHT 92 1442 A 1246 5 562 VPDEENILNELNDPLREEINVFNCRKLVATMP LFANAPPNFVTAMLSKLRFEVFOPGDYIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGG YFGEICLLTKGRTTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDGLDRIGKKN SLILOKFQOLLTKGRTATSVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRDRIGKKN SLILOKFQOLLTKGRTATSVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRDRIGKKN SLILOKFQOLLTHGRATSVSTTSTGAELKEEL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIFERRKDSTSTFNLIPSSMMTRRNTYVCT	00	1420		1202		742	
LVVVAPRFVDADNILINPDTLSLLIAENKTVV APMLDSRAAYSNFWCGMTSQGYYKRTPAYI PIRKRDRRGCFA VPWVHSTFLIDLRKAASRNI. VAFYPPHPDYTWSFDDIIVPASCKQVAEVQMY VCNKEEYGFL PVPIRAHSTLQDEAESFMHVQ LEVMVPSSPSSAQSMAVVSADHIGLVISYL PIRKRDRRGCFA VPHRAHSTLQDEAESFMHVQ LEVMVPSSPSSAQSMAVVSADHIGLVISYL VARVENDEAESFMENVQ LEVMVPSSPSSAQSMAVVSADHIGLVISYL VARVENDEAESFMENVQ LEVMVPSSPSSAQSMAVVSADHIGLVISYL VARVENDEAESFMENVADGE VARVENDEAESFMENVASQAYNLT SALTPILTRSRVLNEDEALTAGRSRAPANLSD VVQLIFUDSNPPFGISTRVSTVLSTK VASMAF QTQAGAQIPIERLASERAITVKVPNNSDWAAR GHRSSANSVVQPQAFVGAVVILDSSNPAAV LHQLNYTLLDGRYSSEPEPYLAVYLHSEPR PNEHNCSASRRIRPESLQGADHRPYTFPISPGT RDPVGSYRLNLSSHPRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHRFVFPPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGGTTAHVGIM LYGVDSRSGHHLDDGDASRGRAIP FCGQRGRFKYEILVKTGWGRGGTTAHVGIM LYGVDSRSGHHLDDGDASRGRAIP FCGQRGRFKYEILVKTGWGRGGTTAHVGIM LYGVDSRSGHHLDDGDASRGRAIP FCGQRGRFKYEILVKTGWGRGGTTAHVGIM LYGVDSRSGHHLDDGDASRGRAIP FCGQRGFKYEILVKTGWGRGGTTAHVGIM LYGVDSRSGHHLDDGDASRGRAIP FCGQRGFFTVHINSLDFQLATT HSLGSMWKIRVWHDNKGLSFAWFLQHIIVRD LQTARSTFFLVNDWLSVETEARGGLVEKEVL AASKASRVPIPSVAALLRFRRLLVALDQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPYYLAILFLRFMSRKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLIGKGQHT GKGVHQRLIGKGQHT SVERWKLTDGS AYGKKMYFLQHGVGTKSSKEMKLTDGS AYGKKMYFLQHGVGTKTSSKEMKLTDGS AYGKKMYFLQHGVGTKTSSKEMKLTDGS AYGKKMYFLQHGVGTKTSSKEMKLTDGS AYGKKMYFLQHGVGTKTSSKEMKLTDGS AYGKKMYFLQHGVGTKTSSKEMKLTDGS AYGKKMYFLQHGVGTKTSSKEMKLTDGS AYGKKMYFLQHGVGTKTSSKEMKLTDGS AYGKKMYFLQHGVGTKTSSKEMKLTDGS AYGKKMYFLQHGVGTKTSSKEMKLTDGS AYGKKMYFLQHGVGTKTSSKEMKLTDGS AYGKKMYFLQHGVGTKTSSKEMKLTDGS AYGKKMYFLQHGVGTKTSSKEMKLTDGS AYGKKMYFLQHGVGTKTSSKEMKLTDGS AYGKKMYFLQHGVGTTKSSKEMKLTDGS AYGKKMYFLQHGVGTTKSSKEMKLTDGS AYGKKMYFLQHGVGTTKSSKEMKLTDGS AYGKKMYFLQHGVGTTKSSKEMKLTDGS AYGKKMYFLQHGVGTTKSSKEMKLTDGS AYGKMYFLQHGVGTTKSSKEMKLTDGS AYGKKMYFLQHGVGTTKSSKEMKLTDGS AYGKFTA	69	1439	A	1223	1	743	
APMLDSRAAYSNFWCGMTSQGYYKRTPAYI PIRKRDRRGCFAVPMVHSTFLIDLRKAASRNL VAFYPHPDVTWSTFDDIIVFAFSCKQAEVQMY VCNKEEYGFLPVPLRAHSTLQDEAESFMHVQ LEVMVPSSPSSAQAWVSADHIGLVISYL 90 1440 A 1227 2 349 NKTSFIFYLKNIVVADLIMITLFFPRIVHDAGF GPWDFKFILCRYTSVLFYAMMOTSIVVLGLITI YDRY/WKVVRLIWDSWMTGIVSFTRAYLLG LGARLVWFGKLILAKGGHGGISWL 91 1441 A 1245 3 1937 LGSSDVRAPQRSELGAESPSRMVASQAYNLT SALTPILITRSRVLNEEPLTLAGFSRAPANLSD VVQLIFLVDSNPFPFGYISNYTVSTKVASMAF QTQAGAQIPIERLASERAITVKVPNNSDWAAR GHRSSANSVVQPQAFVGAVVTLDSSNPAAV LHLQLNYTLLDGRYLSEEPEPYLAVYLHSEPR PMEHNCSASRRIRPESJQGADHRPYTFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTARGTSLFVPBHAVDNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRKYEILVKTGWGRGSTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATT HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSAALLKFRKLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWGAVGADSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGGSTGNGAWASAPCPEPPLSSAAAR GKGVPGRLGKGGVTTKSKERMELTDGS VFGEICLITKGRRFASVRADTYCRLYSLSVD NFNEVLEEVPMMRRAFETVAIDRIGKKN SILLQKFQKDLNTOVPNNQEMELKQIVKH 93 1443 A 1249 180 901 TVPPPFGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERKDSTSTFNILPPSMMTRRNTYVCT							
PIRREDREGCFAVPMYHISTFLÜLIRKAASRNI. VAFYPPHPDYTWSFDDIIVFAFSCKQIAEVQMY VCNKEEYGFLPVPLRAHSTLQDEAESFMHVQ LEVMVPSSPSSAQSMAVVSADHIGLVISYI. 90 1440 A 1227 2 349 NITSFEYLKNIVAVADUMATLFPEFFIVHDAGF GPWDFKFILCRYTSVLFYANMDTSIVVLGLIT/ YDRYWKVVRHL/WDSWMTGISFTRVYLIG LGARLVWFGKILLAKGGHGGISWI. 91 1441 A 1245 3 1937 LGSSDVRAPQRSELGAESPSRMVASQAYNLT SALTPILTRSRVLNEEPLTLAGFSRAPANLSD VVQLIFL/UDSNPPFGYISNYTVSTKVASMAF QTQAGAQIPIERLASERAITVKVPNNSDWAAR GHRSSANSVVQPQAFQAVVTLDSNPPAV HLQLNYTLLDGRYLSEEPEPYLAVYLHSEPR PNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDPVGSYRINLSSHFRWSALEVSVGLYTSLC QYFSEEDVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEIL/KTGWGRGGTTAHVGIM LYGVDSSGHRHLDGBRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETERANGCLVEKEVL AASKASFRVPTPSAALLKFRRILVABLORGF FDKHIWLSIWDRPPRSCFTRICRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVPYVLAILFLFRMSRSKV GWGWGGGSTGNGAWASAPCPEPPLSSAAAR GKGVPQRLJGKGGHT 92 1442 A 1246 5 562 VFDEENINELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVPQPGDYIREG AVGKKMYFLQHAVGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVADRLDRIGKKN SILLQKFQKDLINTGVFNNQEMEILKQIVKH 93 1443 A 1249 180 901 TVPPPFGGPSPAPLHFKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERKDSTSTENNLPPSMMTRRNTYVCT	1 1			ł		}	
90 1440 A 1227 2 349 NKTSFFYLRNIVVADLIMTLITPFRIVIDAGE GPWDFKFILCRYTSVLFYANMDTSIVVLGLIT/ YDRY/WKVVRHL/WDSWMTGI/SFTRVVLLG GPWDFKFILCRYTSVLFYANMDTSIVVLGLIT/ YDRY/WKVVRHL/WDSWMTGI/SFTRVVLLG LGARLWWFGKLLAKGGHGGISWL 91 1441 A 1245 3 1937 LGSSDVRAPGRSELGAESPSRMVASQAYNLT SALTPILTRSRVLNEEPLITLAGPISRAPANLSD VVQLIFLDSNFPFGYISNYTVSTKVASMAF QTQAGAQIPIERLASERATIVKVPNNSDWAAR GHRSSANSVVQPQAFVGAVVTLDSSNPAAV LHLQLNYTLLDGRYLSEEPEPYLAVYLHSEPR PNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRRVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP PCGQRGRFKYEILVKTGWGRGGGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRYPTPSAAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGFGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKQQHT 92 1442 A 1246 5 562 VPDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVPQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLITKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRILDRIGKKN SILLQKRQKDLTGVFNNQENEILKGIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGGGLPPSSPMVSSAHNPN KAEIPERRKDSTSTYNNLPPSMMTRRNTYVCT							
LEMMYPSSPSSAQSMA_VSADHIGLVISYL))					1	
90 1440 A 1227 2 349 NKTSFIFYLKNIVVADLIMTLTFPFRIVHDAGF GPWDFKFILCRYTSVLFYANMDTSIVVLGLIT/ YDRY/WKVVRHL/WDSWMTGI/SFTRYVLLG LGARL/WFGKLILAKGGHGGISWL 91 1441 A 1245 3 1937 LGSSDVRAPQRSELGAESPSRNVASQAYNLT SALTPILITRSRV.INEEPLTLAGF9SRAPANLSD VVQLIFL/DSNPPFFGYISNYTVSTKVASMAF QTQAGAQIPIERLASERAITVKVFMISDWAAR GHRSSANSVAVQOQAFVGAVVTLDSSNPAV LHLQLNYTLLDGRYLSEEPEPYLAVYLHSEPR PNEHNCSASRRIRRESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEDVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHENSLDIPQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSVAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPYLALLFLFRMSRSKV GWGWGFGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VPDEENILNELNDPLREZIVNFNCRKLVATMP LFANADPRFYTYAMLSKLRFEVFOPGDYIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIJGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHFKRSFTSTGEAELKEERL PGRKASCSTAGSGSGGLPPSSMVSSAHNPN KAEIPERRKDSTSTFNLPPSSMMTRRNTYVCT							
GPWDFKFILCRYTSVLFYANMDTSIVVLGILT/ YDRY/WKVVRHL/WDSWMTGI/SFTRYVLLG LGARLVWFGKLILAKGGHGGISWL 91 1441 A 1245 3 1937 LGSSDVRAPQRSELGAESFSRMVASQAYNLT SALTPILTRSRVLNEEPLTLAGF)SRAPANLSD VVQLIFLVDSNFFPFGYISNYTVSTKVASMAF QTQAGAQPIPERLASERAITVKVPNNSDWAAR GHRSSANSVVQPQAFVGAVVTLDSSNPAAV LHLQLNYTLLDGRYLSEEPEPYLAVYLHSEPR PNEHNCSASRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHRFVVPPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCQQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGYDSRSGHRHLDGDRAFHRNSLDIPGIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSVAALLRFRRLLVAELQRGF FPKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCFEPPLSSAAAR GKGYHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFYTAMLSKLRFEVFQPGDYIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSGGLPPSSPMVSSAHNPN KAEIPERKKDSTSTPNNLPPSMMTRNNTYVCT	90	1440	Λ	1227		240	LEVMVPSSPSSAQSMAVVSADHIGLVISYL
91 1441 A 1245 3 1937 LGSBÜVRAPQRSELGAESPSRMVASQAYNLT SALTPILTRSRVLNEEPLTLAGFISRAPANLSD VVQLIFLVDSNIPFPFGYISNYTVSTKVASMAF QTQAGAQIPIERLASERATIVKVPNNSDWAAR GHRSSANSVIVQPQAFVGAVVTLDSSNPAAV LHLQLNYTLLDGRYLSEEPEPYLAVYLHSEPR PNEHNCSASRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASSGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSNAALLRFRRLLVAELQRGF FDKHIWLSJWDRPPRSCTTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGFGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VPDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPHYTAMLSKLRFEVPQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEPYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	1 30 1	1440	A	1227	2	349	
STATE STAT							
91 1441 A 1245 3 1937 LGSSDVRAPQRSELGAESPSRMVASQAYNLT SALTPILTRSRVLNEEPLTLAGFSRAPANLSD VVQLIFL VDSNPFPFGYISNYTVSTK VASMAF QTQAGAQIPIERLASERAITVK VPNNSDW AAR GHRSSANSVVQPQAFVGAVVTLDSSNPAAV LHLQLNYTLLDGRYLSEEPEPYLAVYLHSEPR PNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHDLGDRAFHRNSLDIFQIATP HSLOSMWKIR VWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSVAALLRFRRLL VAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTSINGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFFIEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNOENEILKQIVKH 93 1443 A 1249 180 901 TVPPPFGGPSPAPLHFKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT]]						
SALTPILTRSRVLNEEPLTLAGFSRAPANLSD VVQLIFLVDSNPPFGYISNTTVSTKVASMAF QTQAGAQIPIERLASERAITVKVPNNSDWAAR GHRSSANSVVQPQAFVGAVVTLDSSNPAAV LHLQLNYTLLDGRYLSEEPEPYLAVYLHSEPR PNEHNCSASRRIRESLQGADHRPYTFPISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP PCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIR VWHDNKGLSPAWFLQHHIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPS\AALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCTTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKQQHT 92 1442 A 1246 5 562 VPDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEVPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHFKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	01	1441	Α	1245	2	1027	
VVQLIFLVDSNPFPFGYISNYTVSTKVASMAF QTQAGAQIPIERLASERAITVKVPNNSDWAAR GHRSANSVIVQPQAFVGAVVTLDSSNPAAV LHLQLNYTLLDGRYLSEPEPYLAVYLHSEPR PNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRRVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPFGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	''	1441	Α.	1243	3	1937	
QTQAGAQIPIERLASERAITVKVPNNSDWAAR GHRSSANSVVQPQAFVGAVVTLDSSNPAAV LHLQLNYTLLDGRYLSEEPEPYLAVYLHSEPR PNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVND LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSVAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCKLVATMP LFANADPNTVTAMLSKLRFEVFQPGDYIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	1		-				
GHRSSANSVVQPQAFVGAVVTLDSSNPAAV LHLQLNYTLLDGRYLSEEPEPYLAVYLHSEPR PNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFILVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSVAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPYVLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	1 1						
LHLQLNYTLLDGRYLSEEPEPYLAVYLHSEPR PNEHNCSASRRIRPESLQGADHRPYTIFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPS\AALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 73 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGAELKEERL PGRKASCSTAGSGSGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT							
PNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRPVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIR VWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPS\AALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	1 1						
RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSVAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT		l	l	1			
QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPS\AALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	1 1						
HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPS\AALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	i l	j	ļ	,]			
TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	1 !		l				
FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPS\AALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT				ļ			
LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPS\AALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAH\NPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT				1			
HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPS\AALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	1	l	ľ	1		l	
LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPS\AALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	1	Í		Ì			
AASKASFRVPTPS\AALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEVPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT						ļ	` '
FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT		ţ		ļ			
FLGANAVWYGAVGDSAYSTĞRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT							
DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT		1		ľ		ľ	
GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT		1	1	1	ł	İ	· · · · · · · · · · · · · · · · · · ·
92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT				İ	1		The state of the s
92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	1 1		1	1			
LFANADPNFVTAMLSKLRFEVFQPGDYIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	92	1442	A	1246	5	562	
AVGKKMYFIQHGVAGVITKSSKÈMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT		}		į	ì		
93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT] [ļ	[ĺ	ĺ		
93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	[1	ſ	ſ			,
93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	[]	}	į				1
93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAH\NP\ KAE\PERKDSTSTP\NL\PPSMMTRR\TY\CT			j	ŀ		1	
PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	93	1443	A	1249	180	901	
KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT		ľ				-	
		- 1	ſ	[•	ĺ	
						ŀ	

CEO IN	L CEC TE	T 3/-4	TCC	Dec 1	T	T
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ļ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1	ł	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		1	l	residue of	sequence	
ł	1	1	1		sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
j	1		l	peptide		/=possible nucleotide deletion, \=possible
	<u> </u>	<u> </u>	ļ <u>.</u>	sequence		nucleotide insertion
1	ł	l .	1	1		LAPPSGERSRLARGSTIRSTFHGGQVRDRRAG
J	}		J]	J	GGGGGGVQNGPPASPTLAHEAAPLPAGRPRP
1	1		_			TTNLFTKLTSKLTRRVADEPERIGGPEVTRRP
1	İ		1			ROEDHLSPGGRGCSEL
94	1444	A	1261	3	385	
' '	1444	1 ^	1201	ر ا	303	KFSQWGLTKPKLSNASP/WISLVKKLMKKWS
{	ł	ł	l	ł	l	VTQNLTFREQLEAGIRYFDLRVSSKPGDADQ
J	}	ļ]		1	EIYFIHGLFGIKVWDGLMEIDSFLTQHPQEIIFL
-	l	1		Į.	ĺ	DFNHFYAMDETHHKCLVLRIQEAFGNKLCPA
ł	ł	1	}	ł		CR
95	1445	A	1282	2	550	GPRDNPG\EDPRFEIVEHFGIAWFTFELVARFA
		1		i	***	
1	1	[i	ł	1	VAPDFLKFFKNALNLIDLMSIVPFYITLVVNL
}]	l	j]	VVESTPTLANLGRVAQVLRLMRIFRILKLARH
	1	ŀ	1			STGLRSLGATLKYSYKEVGLLLLYLSVGISIFS
1		1	1	Ì	ł	VVAYTIEKEEN\EGLATIPACWWWATVSMTT
L				l	}	VGYGDVVPGTTAGKLTASACILA
96	1446	A	1294	1	1456	QLLPPSNRENAGLLVGRCLCSAALRPVGDLIT
ł	i	ł		_		SSGQVAVRNAPOAGSAKAGKGKFODNFEFIO
}		ļ	!		j	
			[ĺ	YFKKFFDANCNEKDYNPVAAGQGQETEVAP
ł	ł	ł	}		Ì	SIVAPVLNKPNQCPEGYICVKAGRNPNYGYT
						SFDTFSWAFLSLFRLMTQDYWENLYQLTLRA
1		ĺ	1			AETTYMIF/LV/LVILLGSLYLVTLILAV/VAMA
	ļ	ļ	j !			YEEQNQATLEEAEQKEAEFQQMLEOLKKOO
						EAAQQAATATASEHSREPSAAGRLSDSSSEAS
ĺ	ł	ĺ	i I			KLSSKSAKERRNRRKKRKQKEQSGGEEKDED
1	l	ł				
	l					EFQKSESEDSIRRKGFRFSIEGNRLTYEKRYSS
1						
1		1				PHQSLLSIRGSLFSPRRNSRTSLFSFRGRAKDV
		}	1			GSENDFADDEHSTFEDNESRRDSLFVPRRHGE
						GSENDFADDEHSTFEDNESRRDSLFVPRRHGE
						GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC
						GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD
					·	GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR
. 97	1447	A -	1205	2	2057	GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE
97	1447	A	1295	2	2057	GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE
97	1447	A	1295	2	2057	GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI
97	1447	A	1295	2	2057	GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE
97	1447	A	1295	2	2057	GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE IQTQLPTKSSQQLRKGGNCVRCKMQMNFLAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEOQGALS
97	1447	A	1295	2	2057	GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE IQTQLPTKSSQQLRKGGNCVRCKMQMNFLAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEOQGALS
97	1447	A	1295	2	2057	GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP
97	1447	A	1295	2	2057	GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE IQTQLPTKSSQQLRKGGNCVRCKMQMNFLAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEEELSVLESCLHRWRTEIENDTR
97	1447	A	1295	2	2057	GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV
97	1447	A	1295	2	2057	GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI
97	1447	A	1295	2	2057	GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN
97	1447	A	1295	2	2057	GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQE\DLIKTGQPLVGIETLPPDLRDFV
97	1447	A	1295	2	2057	GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLK YRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWA YIFDHRESRWMK YNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS
97	1447	A	1295	2	2057	GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEVDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEOPSRSDFSK
97	1447	A	1295	2	2057	GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEVDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEOPSRSDFSK
97	1447	A	1295	2	2057	GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLK YRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWA YIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE
97	1447	A	1295	2	2057	GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE IQTQLPTKSSQQLRKGGNCVRCKMQMNFLAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEVDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ
97	1447	A	1295	2	2057	GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE IQTQLPTKSSQQLRKGGNCVRCKMQMNFLAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEVDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA
97	1447	A	1295	2	2057	GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE IQTQLPTKSSQQLRKGGNCVRCKMQMNFLAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT
97	1447	A	1295	2	2057	GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEVDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL
97	1447	A	1295	2	2057	GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEVDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL
97	1447	A	1295	2	2057	GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE IQTQLPTKSSQQLRKGGNCVRCKMQMNFLAE EVLLKYRITFYNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQE'DLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF
97	1447	A	1295	2	2057	GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE QRAMSIASILTNTVE OTOLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLK YRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE
97	1447	A	1295	2	2057	GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLK YRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWA YIFDHRESRWMK YNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT
97	1447	A	1295	2	2057	GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLK YRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWA YIFDHRESRWMK YNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER
						GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEVDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLIKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCËR FARIMLSLSRTPADGR
97	1447	A	1295	2	2057	GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEVDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR
						GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE IQTQLPTKSSQQLRKGGNCVRCKMQMNFLAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEVDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV
						GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE IQTQLPTKSSQQLRKGGNCVRCKMQMNFLAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEVDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV
						GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEVDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD
	1448	Ä	1304	118	453	GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE OTOLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLK YRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWA YIFDHRESRWMK YNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEVDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS
98						GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEVDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	}	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
ľ			ĺ	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			ĺ	peptide	Ţ	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
i						DSRLLLFDAPDLRLSPPSGALLQVLDLRDPQF
					i	SATPVLASDVIHAQSRDLPRIFRVTTSQLAVPP
			!			TTCTVLLLAESEGERERWLQVLGELQRLLLD
		Ì				ARPRPRPVYTLKEAYDNGLPLLPHTLCAAILD
1		ļ				QDRLALGTEEGLFVIHLRSNDIFQVGECRRVQ
1 1				1	[QLTLSPSAGLLVVLCGRGPSVRLFALAELENI
						EVEVPKIPESRGCQVLAAGSILQARTPVLCVA
ŀ						VKRQVLCYQLGPGPGPWQRRIRELQAPATVQ
!						SLGLLGDRLCVGAAGGFALYPLLNEAAPLAL
						GAGLVPEELPPSRGGLGEALGAVELSLSEFLL
1 1						LFTTAGIYVDGAGRKSRGHELLWPAAPMGW
ŀ						GYAAPYLTVFSENSIDVFDVRRAEWVQTVPL
1						KK\VRPLNPEGSLFLYGTEKVRLTYLRNQLAE KDEFDIPDLTDNSRRQLFRTKSKRRFFFRVSE
}			1			EQQKQQRREMLKDPFVRSKLISPPTNFNHLV
1						HVGPANGRPGARDKSP
100	1450	Α	1318	918	190	SLCVPGPVDTGTFAVMSVMVGSVTESLAPOA
100	1450	2 k	1310	710	150	LNDSMINETARDAARVQVASTLSVLVGLFQV
1						GLGLIHFGFVVTYLSEPLVRGYTTAAAVQVF
						VSQLKYVFGLHLSSHSGPLSLIYTVLEVCWKL
1						PQSKVGTVVTAAVAGVVLVVVKLLNDKLQQ
1						QLPMPIPGELLTLIGATGISYGMGLKHRFEAG
						PPVAPNTQLFSKLVGSAFTIAVVGFAIAISLGK
						IFALRHGYRVDSNQVWVMRDV
101	1451	A	1353	220	445	DWPDLFTYPLIGSPKCFQSARPE\RMYRRTVR
						SSHGNHALQEVLPRSGHGTEFTKQKHLEAAD
1						HGHPPARMSIFSR
102	1452	A	1363	542	2	AHLLMLNLAL\TDLL\YLTSLPFLIHYYASGEN
1 1			' '			WIFGDFMCKFIRFSFHFNLYSSILFLTCFSIFRY
						CVIIHPMSCFSIHKTRCAVVACAVVWIISLVA
1 1		ľ			ı	VIPMTFLITSTNRTNRSACLDLTSSDELNTIKW
						YNLILTA\LLCLPLVIVTLCYTTIIHTLTHGHAN
						\DSCLKQKARRLTILLL
103	1453	Α	1371	2	410	CHSTESSSDFILPGDYLLGGLCPLHSGCLQV\C
					1	SFNEHGYHLFQAMRLAVEEINNSTALLPNITL
1		l		-		GYQLYDVCSDSANVYATLRVLSLPGQHHIEL
			ļ			QGDLLHYSPTVLAVIGPDSTNRAATTAALLSP
104	1454		1277		422	FLVPMLLEQ
104	1454	Α	1376	3	432	NSRVEDRS/NMSLWTQNITVCPVRNVTRDGG
						FGPWSPWQPCEHLDGDNSGSCLCRARSCDSP
}	İ	}	1	ļ		RPRCGGLDCLGPAIHIANCSRNGAWTPWSSW
.		- 1	ļ		Į.	ALCSTSCGIGFQVRQRSCSNPAPRHGGRICVG
105	1455	Ā	1379	2	396	KSREERFCNENTPCPVPIF
103	1433	^	13/7	-	350	GLGLLYLIFAAVEGVMRVIGGSNHLAVVLDD
		1				IILAVIDSIFVWFIFISLAQTMKTLRLRKNTVKF
		J	ļ	J		SLYRHFKNTLIFAVLASIVFMGWTTKTFRIAK CQSDWMERWVDDAFWSFLF\SLILIVIMFLW
		1	ļ	ļ		RPSA
106	1456	A	1383	1	432	EDGHGGWSSRCLVDHAEEGHREPWKRLCIW
100	1730	^	1202	•	47€	QRGGHEIRFAFYFPGHPLLSPQICLAPETPPRG
		ĺ	ļ			CPPVSSLHFISLO/RLPRDCQELFQVGEROSGL
ļ	j					FEIQPQGSPPFLVNCKMTSGTFWTCRTDSRVF
		[[QNANPSNAAHSEDQPTP
107	1457	A	1386	719	558	FFFVTRSHSVAQAECSGVFTAHRSLDLVGSSN
"	1.0,	·^	1300	,,,	230	YPALSLQSSWDHRHTWLIFAFL
108	1458	Ā	1397	631	2	RVAISLLCAAIFISFMVQSAGKRWPTGVMLM
	1.50		.371		-	VVVLFAFLYSWPIQALLPTYLKTDLAYNPHT
		- 1	1			VANVLSFSGFGAAVGCCV/GGFLGDWLGTRK
		ļ	- 1	•		AYVCSLLASQLLIIPVFAIGGANVWVLGLLLF
						THE COLUMN TO TH

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1100	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	!	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	i	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	uciicc		914	ng to first	acid residue	
dence	j	}	1714	amino acid	of peptide	Q=Glutamine, R=Arginine, S=Serine,
1		J		residue of	sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
	İ	į.	ì	peptide	sequence	1-1 yrosine, X=Unknown, *=Stop codon,
		1		sequence		/=possible nucleotide deletion, \=possible nucleotide insertion
	 	 	 	sequence	 	
	ļ		1			FQQMLGQGIAGILPKLIGGYFDTDQRAAGLG
ļ	1	1		ļ	ļ	FTYNVGALGGALAPIIGALIAQRLDLGTALAS
109	1459	A	1402	15	207	LSFSLTFVVILRNRRPGKSLVR
109	1439	A	1402	13	387	VLVALPDT\VTSETVVTEVLGHRVTLPCLYSS
	ŀ			ļ		WSHNSNSMCWGKDQCPYSGCKEALIRTDGM
		(ľ	RVTSRKSAKYRLQGTIPRGDVSLTILNPSESDS
110	1460	 	1401	<u> </u>	1200	GVYCCRIEVPGWFNDVKINVRLNLQRASTT
110	1460	Α	1421	3	350	HEDLSSLLTRGSGNQERERQLKKLISLRDWM
	ľ			ļ		LAELAFPVGVLATCA*SLLSC*YCVILFPCSCF
1	}					FFHSPDALFSLLLLSCYFPSYCFFYYLFFSSSPL
	1	ļ				CLLLASSPFPLFILLASL
111	1461	A	1426	2	344	FTSTMTKPFEKESEQPA*ATLAFGAQTSTTAD
	'	l				QCALKPDLSYLNNSSSSSSTPATSAGGGIFGSS
1	ţ	l			1	TSSSNPPVATFVFGQSSDPVSSYGFVNTAESST
						SDSLLFSQDSKLATTS
112	1462	A	. 1434	46	372	TTSWTTSCTRSCT*SGASSGPGWTPRTTWWR
						SRRSSQRTCSRACSGAWSRTW*RSS*TSSSSC
1						STSCSSSSSRSCGRPGGPLGARGVHITSCLNSC
		<u> </u>				MSSSTTSSTTSTF
113	1463	Α	1439	3	292	HEDIMTHYDRLVDE*ALNAGKQRYEKMISG
	ŀ					MYLGEIVRNILIDFTKKGFLLRGQISEMLKTR
L					j	GIFLTFLLSNFLIVCVLLFYVSFYLFQSCINFVL
114	1464	Α	1463	1	396	KQQAVPEPHSSTTTPQEQEQNWYGQDLLNLQ
						QRTKVHLPGHKTGPAVAKDTPEPVKKEFTVP
1	[ſ				ATSQGP*SPFSEEPPLPPSNEEVPPTLPP*EPQS
1	1	1			1	EDP*KNA*LKQMHAATTHWQQHQQHQVGC
<u> </u>						QYHGIMQ
115	1465	A	1464	291	2	AGSYPSMVWSCHWGVTQKRRAL*VYSFEEG
Į.		ļ	1			GRRKCGQYWPLEKDSRIRFGFLTVSNLGVEN
1		Ì			ļ	MNHYKKSTLEILNPEVNPGFFFLTLWKQGEN
		<u> </u>				NYCN
116	1466	A	1465	667	337	LPPQRPA*TDSYSTCNVSSGFLAGQSHNIHLQ
		ĺ	[YWTKYQVWEWLQHFLDTNQLDANCIPFQEF
	1		í i		1	DINGEHLCSMSLQEFTRAAGTAGQLLYSNLQ
<u> </u>					L i	HLKWNGDSLFLCLSLPC
117	1467	Α	1479	1	381	GTSGGPKRVLVTERFPWQNPLPVNRGQAQR
1			1			VLGPSNSFQRVPLQAQKLVSSHKPGQNQKHK
]	j l		. 1	QLQATSVPHPVCMPLNNTQKSKQPLPSAPEN
						NPEEELASDPNNEESL*RPWALEDFEIGRPLG
			<u> </u>		,	KGK
118	1468	Α	1485	3	385	TYLWL*GNPPFYEKNDGGLFELILRAKDEFNS
				i		PYWDDMSDSAKHFIRPLTGRDP*KPFPCDQPL
[QHPWIEGHTCLDNNIHQAASEPINNNFAESKR
1					ļ	NLAFLATGVVRHMRKLFMGANLEGPGPTVS
						H
119	1469	Α	1486	1	398 .	GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL
			()	ľ		NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE
				l		KALTKFLKWVNWDLPQEAKQALELLGKWK
]		PMDVKDSLELLSSHYTNPTVRRYAVARLRQA
			_			DDEDLLMYL
120	1470	Α	1497	3	999	MGESPAV*GYFVLAGMNSAGLSFGGGAGKY
]]		'		ļ		LAEWMVHGYPSENVWELDLKRFGALQSSRT
						FLRHRVMEVMPLMYDLKVPHWDFQTGRQL
						RTSPLYDRLDAQGARWMEKHGFERPKYFVP
1	.	Ì				PDKDLLALEQSKTFYKPDWFDIVESEVKCCK
1 1	1			ļ	,	EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS
						NDLDVPVGHIVHTGMLNEGGGYENDCSIARL
1	į			J	ļ	NKRSFFMISPTDQQVHCWAWLKKHMPKDSN
				ì		LLLEDVTWKYTALNLIGPRAVDVLSELSYAP

SEQ ID NO: of nucl- eotide	SEQ ID NO: of peptide seq-	Met hod	SEQ ID NO: in USSN	Predicted beginning nucleotide location	Predicted end nucleotide location corresponding	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine.
seq- uence	uence		09/496 914	correspondi ng to first amino acid residue of	to last amino acid residue of peptide sequence	M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide sequence		/=possible nucleotide deletion, \=possible nucleotide insertion
						MTPDHFPSLFCKEMSVGYANGIRVMSMTHT GEPGFMLYIPIEYRWGFTMLSTLVSNS
121	1471	A	1498	3	306	AQFLLVGWDHIL*LIVL*TNLTELGRTTCDQN WPNSPDVLNHGCFYMQCLSKDCTIGYVSRE MLVAHTHTVEEHTGTHLQYVSWPDHSVPDD SSDFVEFEN
122	1472	A	1533	121	329	LGLFSFVWTEVLEEPKDFSCETEDFKTLHCT WDPGTDTALGWSKQPSQSYTLFES*VGSGYII DNFFLA
123	1473	A	1547	111	408	DARTTWKPRNGSSGIWPGDGAK*PPAVEQAE RGHVEMIEKLTFLNLHTSEKDKGGNTALHLA AKHGHSPAVQVLLAQWQDINEMNEKQQTPL HVAADRG
124	1474	A	1555		745	MTFDDDDKNTYGVALVWKKFQTQSLRLSDL HRKSHLWRGIVSITLIEGRDLKAMDSNGLSDP YVKFRLGHQKYKSKIMPKTLNPQWREQFDF HLYEERGGVIDITAWDKDAGKRDDFIGRCQV DLSALSREQTHKLELQLEEGEGHLVLLVTLT ASATVSISDLSVNSLEDQKEREILKRYSPLRI FHNLKDVGFLQVKVIRAEGLMAADVTGKSD PFCVVELNNDRLLTHTVYKNLNPEWNKVFTL *VALVWKKFQTQSLRLSDLHRKSHLWRGIVS ITLIEGRDLKAMDSNGLSDPYVKFRLGHQKY KSKIMPKTLNPQWREQFDFHLYEERGGVIDIT AWDKDAGKRDDFIGRCQVDLSALSREQTHK LELQLEEGEGHLVLLVTLTASATVSISDLSVN SLEDQKEREEILKRYSPLRIFHNLKDVGFLQV KVIRAEGLMAADVTGKSDPFCVVELNNDRLL THTVYKNLNPEWNKVFTL
125	1475	A	1556	57	509	GGPAPNSRYAEP*KNSLAMT*AHADCENYVA CGGLDNICSIYNLKTREGNVRVSRELPGHTGY LSCCRFLDDSQIVTSSGDTTCALWDIETAQQT TTFTGHSGDVMSLSLSPDMRTFVSGACDASS KLWDIRDGMCRQSFTGHVSDINAVS
126	1476	A	1592	3	178	KSEKSCVSSLAHFGTSCQRDYDAMVKLVETL EMLPTCDLADOHNIKFHYAFALNR*ER
127	1477	A	1612	1	497	TESPLLVRPYLPYITKSELHAIMTAGFSTIAGS VLGAYISFGVPSSHLLTASVMSAPASLAAAKL FWPETEKPKITLKNAMKMESGDSGNLL*AAT QGASSSISLVANIAVNLIAFLALLSFMNSALA WVGNMFDYPQLSFELICSYIFMPFSFMMGVE WPDSFM
128	1478	A	1619	286	486	CCMNSKAQESVFKNVLCNPPALSEMPDVKA EDEVDFRASSISEEVAVGSIAATLKMKQGPM TOAINR
129	1479	A	1627	1	395	PTRGALRYWIFGRFLCNIWAAVDVRCCTATI MGLCIISIDRYVGVSYPLRYPTIVTQRRGLMA LLCVWALSLVIYIGPLLGWRHPAPEDETICQI NEEPGYVLFSTPGSFYLPLAIMLVMN*RVYRV AKTE
130	1480	A	1638	2	466	DPRVRTKIVNRKTTIYEIQDKTGSMAVVGKG ECHNIPCEKGDKLRLFCFRLRKRENMSKLMS EMHSFIQIQKNTNQRSHDSRSMALPQEQSQHP KPSEASTTLPESHLKTPQMPPTTPSSSSFTKVT KDKDIK*LLFNLYSSVEILPEVLHLKT
131	1481	A	1651	607	3	LAEGGDVFDCVLNGGPLPESRAKALFRQMVE AIRYCHGCGVAHRDLKCENALLQGFNLKLTD FGFAKVLPKSHRELSQTFCGSTAYAAPEVLQ GIPHDSKKGDVWSMGVVLYVMLCASLPFDD

COROLID	TOPOTO	1 1 6-4	TOPO	I No. Proc. 1		
SEQ ID NO: of	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	Ì	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uenœ			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1	1	Į	l	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		ĺ		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	ļ]	1	peptide)	/=possible nucleotide deletion, \=possible
L			L	sequence		nucleotide insertion
1		}				TDIPKMLWQQQKGVSFPTHLSISADCQDLLK
- }				[RLLEPDMILRPSIEEVSWHPWLAST**KQWQV
L		<u></u>	<u> </u>			LSNKVGGESKPKKKK
132	1482	Α	1656	150	48	LVAKSLLYCGCLFFLLQLAKNVGNNSFNDIM
				Į		EANLTSPSPKPTPSSDM*VFLIY*TYFGAWHV
· L	1	l				VDAQ
133	1483	A	1660	3	406	RKHIKLLIQKLSDVP*ECQNNQL*KLTEICEKE
1	1	1		1	}	KKEFKKKMDDQRPEKITEA*SKDKSPMEEEK
-				1		TEMIRSYIQEVGRYIKRLEEAQSKRLEKLREK
i		1			}	HKEIRQPILDEKPKGEGSSSFLSETCHEDTSWF
	1			!		PNFTP
134	1484	A	1666	1276	466	PGSTHASARITIY*L*IILSNATEVDNNFSKPPP
		1	1 200	1		FFPAGAPPASSSSSSSSSPPTVSTAPPLIPPPGF
		1		1		PPPPGAPPPSLIPTIESGHSSGYDSRSARAFPYG
1	1	1		}		
	}					NVAFPHLPGSAPSWPSLVDTSKQWDYYARSS
1	1	1		ļ		SSSSSSSSSSSSPRDRDRER*RTREREREDHS
1						PTPSVFNSDEERYRYREYAERGYERHRASRE
	1	ļ				KEERHRERRHREKEETRHKSSRSNSRRRHESE
ŀ	i	Ì		İ		EGDSHRRHKHKKSKRSKEGKEAGSEPAPEQE
L		ļ <u> </u>				STEATPAE
135	1485	Α	1673	1	417	PTRPVNSSQAFALVYYTLGALGGNLIAHMGL
1					'	GYRYWAGIGVLQSCESALTHYRLVANHVAS
	1					DISLTGGSVVQRIRLPDEVENPGMNSGMLQE
	i					DLIQYYQFLAEKGDVQAQVGLGQLHLHGGR
L	<u> </u>	<u> </u>				GV*QNHQRAFDYFNLAA
136	1486	Α	1678	525	9	ANTSLSSAAVSAVSPPPCRTSTATTLPPPMPSF
1	ì	l				FCVFPSPSMSPSPSEFLSCIASVSRVHSLSSSSS
	1			!		GSSSTASSLNFSAIMGSSSATASWVLSTASTPP
1	}	ļ				CPSALPSSPAQES*SLAASSSAWPVAGISPSGA
1	1	ļ	ł			CTFPAGSASGAAKAPSPSWRCPSFRALFSLLD
						SSSLSL
137	1487	A	1680	1	2999	AHRDEIQRKFDALRNSCTVITDLEEQLNQLTE
i						DNAELNNONFYLSKOLDEASGANDEIVOLRS
ľ	1			ĺ		EVDHLRREITEREMOLTSOKOTMEALKTTCT
	ŀ	1				MLEEQVMDLEALNDELLEKERQWEAWRSVL
ł	1	1	l			GDEKSQFECRVRELORMLDTEKOSRARADO
]		İ			RITESRQVVELAVKEHKAEILALQQALKEQK
J	1	ļ]			LKAESLSDKLNDLEKKHAMLEMNARSLOOK
			1	İ		LETERELKQRLLEEQAKLQQQMDLQKNHIFR
					'	LTQGLQEALDRADLLKTERSDLEYQLENIOV
		1	[LYSHEKYKMEGTISQQTKLIDFLQAKMDQPA
						KKKKVPLQYNELKLALEKEKARCAELEEALO
1	1		1			KTRIELRSAREEAAHRKATDHPHPSTPATARO
						QIAMSAIVRSPEHQPSAMSLLAPPSSRRKESST
1			1 :			
1						PEEFSRRLKERMHHNIPHRFNVGLNMRATKC
1	J	}				AVCLDTVHFGRQASKCLECQVMCHPKCSTC
1					İ	LPATCGLPAEYVTHFTEAFCRDKMNSPGLQT
	ı					KEPSSSLHLEGWMKVPRNNKRGQQGWDRK
						YIVLEGSKVLIYDNEAREAGQRPVEEFELCLP
					l	
						DGDVSIHGAVGASELANTAKADVPYILKMES
					ļ	DGDVSIHGAVGASELANTAKADVPYILKMES HPHTTCWPGRTLYLLAPSFPDKQRWVTALES
					ı	DGDVSIHGAVGASELANTAKADVPYILKMES HPHTTCWPGRTLYLLAPSFPDKQRWVTALES VVAGGRVSREKAEADAKLLGNSLLKLEGDD
					į	DGDVSIHGAVGASELANTAKADVPYILKMES HPHTTCWPGRTLYLLAPSFPDKQRWVTALES VVAGGRVSREKAEADAKLLGNSLLKLEGDD RLDMNCTLPFSDQVVLVGTEEGLYALNVLK
						DGDVSIHGAVGASELANTAKADVPYILKMES HPHTTCWPGRTLYLLAPSFPDKQRWVTALES VVAGGRVSREKAEADAKLLGNSLLKLEGDD
						DGDVSIHGAVGASELANTAKADVPYILKMES HPHTTCWPGRTLYLLAPSFPDKQRWVTALES VVAGGRVSREKAEADAKLLGNSLLKLEGDD RLDMNCTLPFSDQVVLVGTEEGLYALNVLK
						DGDVSIHGAVGASELANTAKADVPYILKMES HPHTTCWPGRTLYLLAPSFPDKQRWVTALES VVAGGRVSREKAEADAKLLGNSLLKLEGDD RLDMNCTLPFSDQVVLVGTEEGLYALNVLK NSLTHVPGIGAVFQIYIIKDLEKLLMIAGEERA
						DGDVSIHGAVGASELANTAKADVPYILKMES HPHTTCWPGRTLYLLAPSFPDKQRWVTALES VVAGGRVSREKAEADAKLLGNSLLKLEGDD RLDMNCTLPFSDQVVLVGTEEGLYALNVLK NSLTHVPGIGAVFQIYIIKDLEKLLMIAGEERA LCLVDVKKVKQSLAQSHLPAQPDISPNIFEAV
						DGDVSIHGAVGASELANTAKADVPYILKMES HPHTTCWPGRTLYLLAPSFPDKQRWVTALES VVAGGRVSREKAEADAKLLGNSLLKLEGDD RLDMNCTLPFSDQVVLVGTEEGLYALNVLK NSLTHVPGIGAVFQIYIIKDLEKLLMIAGEERA LCLVDVKKVKQSLAQSHLPAQPDISPNIFEAV KGCHLFGAGKIENGLCICAAMPSKVVILRYN ENLSKYCIRKEIETSEPCSCIHFTNYSILIGTNK
						DGDVSIHGAVGASELANTAKADVPYILKMES HPHTTCWPGRTLYLLAPSFPDKQRWVTALES VVAGGRVSREKAEADAKLLGNSLLKLEGDD RLDMNCTLPFSDQVVLVGTEEGLYALNVLK NSLTHVPGIGAVFQIYIIKDLEKLLMIAGEERA LCLVDVKKVKQSLAQSHLPAQPDISPNIFEAV KGCHLFGAGKIENGLCICAAMPSKVVILRYN

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ļ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		j	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1 4000	ļ		71.4	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
]])	ļ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	sequence	/=possible nucleotide deletion, \=possible
	Ì	ļ		sequence		nucleotide insertion
——			<u> </u>	sequence		
						YGRRSRTDDLKWSRLPLAFAYREPYLFVTHF
1		1				NSLEVIEIQARSSAGTPARAYLDIPNPRYLGPA
}	ļ	ł			ļ	ISSGAIYLASSYQDKLRVICCKGNLVKESGTE
120	1 100	ļ.,	1606			HHRGPSTSRR*PASPLPQYQGQRAFLQGRRK
138	1488	Α	1686	2	526	GRPQGPAPGAGSPPESGPGLWAALGCSLVWV
	j	i '				PLCCLGGAAGRL*ARSGKSGLRRRRAHAGPP
]	j		J		}	PGGPCNSCP*CSAPESGGRGPLPGPGTGGVCS
		İ				CWTRGCQTTARTAAAAAAPGPAGRRPPGGA
]		İ			İ	PQNGSCAASASQEAAAPPPMCPPGRRWAVAS
		<u></u>				PPETRCPAAPGTRCRRLEAA
139	1489	A	1693	3	376	LPSMSNCTSCFRLQSRTES*IRQAGHLLGRNE
		1			1	FIETKALGCAWFSLCYYLVLYFESSHKVDFVF
				!		IV*CFSTPPGAQMTIMSQACAERCNIMRLVDR
	i					RWAGIAKGVGTQKIIGRVHLGEQKALGL
140	1490	A	1704	3	376	ERTNKFIKELIMDGKNLIAATKSLSVAQRKFA
		ļ				HSLRDFKFEFIGDAVTDDERCIDASLREFSNFL
,		}	J		ŀ	KNLEEQREIMVS*EGCKLISQLSRGKKIWIWK
'		[LVLVEVVKHLSLGTVVHCNGKMRFPEP
141	1491	A	1743	1	362	LITNKVFVARELSCLDVHLDSTGSTAVVADO
				_		DKLELELVLKGSYEDTQTSFLGTASAFRFHY
						MAAL*TELSGRLRSSKSNGWNGDNSTGYLTV
			1			PLRPLTIVKEVTMDVPAPNVRGLNWMG
142	1492	A	1769	7	406	NNPSTLPRGS*PMSPRTTMGRRRQRRREHKSS
i i	ښرا د	**	1,705	•	1 400	LSLASSTVGPGGQIVHTETTEVVLCGDPLSGF
ļ		İ				<u>-</u>
						GLQLQGGIFATETLSSPPLVCFIEPDSPAERCG
]]]		,	LLQVGDRVLSINGIATEDGTMEEANQLLRDA ALAHKVV
143	1493	A	1789	1	447	QMLRNGGDQNTVPDYHFADRIRELL*PTEDQ
143	1423	Α .	1707	. *	, 447	KNCIP*DTYLRPSALGNIVEEVTHPCSPGPCPA
						NELCEVNRKGCTSGDPCLPYFCVQGCKLGQA
. 1		ĺ	[SDFIARQGTLIQVPSSAGEVECYKICSCGQSGL
144	1494		1814		404	LENCMEMHCMDLPTDTSALVR
144	1494	Α	1814	1	404	PGRRFRPRLSQAGTDSGS*VFPDSFPSAPAEPL
						PYFLQEPQDAYIVKNKPVELRCRAFPATQIYF
			1 1	i		KCNGEWVSQNDHVTQEGLDEATGLRVREVH
						IEVSRQQVEELFGLEDYWCQCVAWSSAGTTK
145	1405		1005			SRRAYVRI
145	1495	A	1827	26	448	XVEEKHADTWRSXCLSDFFFHAAKXLCXE*N
				[CGDAISLSVGDHFGKGNGLTWAEKFQCEGSE
		!			ļ	THLALCPIVQHPEDTCIHSREVGVVCSRYTDV
				ĺ		RLVNGKSQCDGQVEINVLGHWGSLCDTHWD
						PEDARVLCRQLNCGTAL
146	1496	Α	1828	574	333	QHEGGDLRRRQLGEIQLTVRYVCLRAASAC*
				j		SMAAET*HHVPASGADPYVRVYLLPERKWA
				}	l	CRKKTSVKRKTLEPLFDET
147	1497	Α	1855	1	372	ERLVLTSEHCLVLTLFWPSWTYHTLLLSRQH
					j	VRRLPKLTHAEHDHLASIMNKLLTNYDNLFE
]				ľ	Ì	TSVTYSMG*HGAPTGSEAGANWNH**LHAH
1		•			ļ	YYPPLLRSDTVRKFMVGSQMLAQAQRDLTPE
1]			Q
148	1498	A	1879	568	7	LLSALDDKGGTQPSASFSNAPTIVCVTACPAG
	/-				<i>'</i>	IAHTYMAAEYLEKAGRKLGVNVYVEKQGAN
	ĺ				ł	GIEGRLTADQLNSATACIFAAEVAIKESERFN
. 1				- 1	ļ	GIPALSVPVAEPIRHAEALMQQALTLKRSDET
]		RTVQQDTQPVKSVKTELKQALLSGISFAVPLI
149	1499	A	1880	611	24	VAGGTQVA*AV*RQGISSLHDVQVRTWNS
147	1477	Λ.	1000	011	24	GLNSENALSNEAMERGWQCLRLFAERLQDIP
· }		ļ]	}	}	PSQIRVVATATLRLAVNAGDFIAKAQEILGCP
		1			l	VQVISGEEEARLIYQGVAHTTGGADQRLVVD

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion IGGASTELVTGTGAQTT*LFSLSMGCVTWLER YFADRNLGQENFDAAQKAAREVLRPVADEL RYHSWKEVRGASVTVQALQEIMMAQGMDE
150	1500	A	1894	2	750	RITMEIWPVD GRVDFFHTDYRPLIRDSNNYVLDEQTQQAPH LMPPPFLVDVDGNPHPTKYQRLVPGRENSAD EHLIPQLGYVATSDGEVIEQIISLQTNDNDERS PESSILDGMIRQLQQQDQRMGADQDTIPRG LSNGEETPRRGFRRLSLDIQSPPNIGLRRSQV EGVRQMHQNAPRSQIATERDLQAWKRRVVV PEVPLGIFRKLEDFRLEKGEEERNLYIIGRKRK TLQLSHKSDSVGLVSQSRPRTCRRKYP
151	1501	Α	1900	141	785	GKTIQIQTTMQNKYKTVQKQYKTIPKNKKA MEMQIKKQFQDTCKVQTKQYKALKNHQLEV TPKNEHKTILKTLKDEQTRKLAILAEQYEQSI NEMMASQALRLDEAQEAECQALRLQLQQEM ELLNAYQSKIKMQTEAQHERELQKLEQRVSL RRAHLEQKIEEELAALQKERSERIKNLLERQE REIETFDMESLRMGFGNLVTLDFPKEDYR
152	1502	A	1915	2	377	LVRLLDTQRDGLQNYEALLGLTNLSGRSDKL RQKIFKERALPDIENYMFENHDQLRQAATEC MCNMVLHKEVQERFLADGNDRLKLVVLLCG EDDDKVQNAAAGALAMLTAAHKKLCLKMT QVTT
153	1503	A	1921	1	237	AYQSLRLEYLQIPPVSRAYTTACVLTSAAVQL ELITPFQLYFIPELIFKHFQIWRLITNFLFFVPFG FNFLLYMIFLYT
154	1504	A	1928	2	354	EMVEGGEGKMCINTEWGGFGDNGCIDDIRTR YDTEVDEGSLNPGKQRYEKMTSGMYLGEIV RQILIDLTKQGLLFRGQISERLRTRGIFETKFLS QIESDRLALLQVRRILQQLGLD
155	1505	A	1929	2	369	TEIAKIKMEAKKKYEKELTMFQNDFEKACQA KSEALVLREKSTLERIHKHQEIETKEIYAQRQ LLLKDMDLLRGREAELKQRVEAFESYQLELK DDYIRTYRLIEDDRINIQISGHWQESP
156	1506	·A	1935	1	270	VTRKLPIFIVDAFTARAFRGSPAADCLLENEL DEDMHQKIAREMNLSETAFIRKLHPTDNFAQ RSCFGLIWFTPTTDLQILTSSILPSIL
157	1507	A	1936	584	305	ESKVNNEKFRTKSPKPAESPQSATKQLDQPTA AYEYYDAGNHWCKDCNTICGTMFDFFTHMH NKKHTQGQFQKSSDFQKEELQQTFLPPERQG
	1508	A	1939	1	423	TTHRLNVTAEPPCTSMPIYWMPDVPHRCTTA NTCPVDLTDYCAQNGFYCLVYGFLPYGSLED RLHCQTQACPPLSWPQRLDILLGTARAIQFLH QDSPSLIHGDIKSSNVLLDERLTPKLGDFGLA RFSRFAGSSPIQSSM
159	1509	A	1974	3	401	HTSTARLLLHRGAGKEAVTSDGYTALHLAAR NGHLATVKLLVEEKADVLARGPLNQTALHL AAAHGHSEVVEELVSADVIDLFDEQGLSALH LAAQGRHAQTVETLLRHGAHINLQSLKFQGG HGPAATLLR
160	1510	A	1982	2	417	KFLKDLEKQYNKEEPHLSEIGSCFLQNQEGFA IYSEYCNNHPGACLELANLMKQGKYRHFFEA CRLLQQMIDIAIDGFLLTPVQKICKYPLQLAEL、 LKYTTQEHGDYSNIKAAYEAMKNVACLINER KRKLESIDKIA
161	1511	A	1984	4	770	RETGSVSLSPSGLEGAESYAVSPILYSSPDVKE LWLETLQGQRHSHTGVKSTPGQSAAILMKLR SSHNASKTLNANNMETLIECQSEGDIKEHPLL

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion ASCESEDSICQLIEVKKRKVLSWPFLMRRLS PASDFSGALETDLKASLFDQPLSIICGDSDTLP RPIQDILTILCLKGPSTEGIFRRAANEKARKEL KEELNSGDAVDLERLPVHLLAVVFKDFLRSIP RKLLSSDLFEEWMGALEMQDEEDRIEALK
162	1512	A	1986	864	501	LLNSGLFSAPDGSNLEMRLTRGGNMCSGRIEI KFQGRWGTVCDDNFNIDHASVICRQLECGSA VSFSGSSNFGEGSGPIWFDDLICNGNESALWN CKHQGWGKHNCDHAEDAGVICSSKD
163	1513	A	2001	419	187	AVDLSIDESSLTGETTPCSKVTAPQPAATNGD LASRSNIAFMGTLVRCGKAKGVVIGTGENSE FGDIINLSTFVVHS
164	1514	A	2012	284	597	SLLCLFPGTSTVVCKPIVIETQLYVIVAQLFGG SHIYKRDSFANKFIKIQAIEILKIRKPNDIETFKI ENNWYFVVADSSKAGFTTIYKWERETGFYSH QSFTR
165	1515	A	2013	2	403	EDPEELGHFYDYPMALFSTFELFLTIIDGPANY NVDLPFMYSITYAAFAIIATLLMLNLLIAMMG DTHWRVAHERDELWRAQIVATTVMLERKLP RCLWPRSGICGREYGLGDRWILRVEDRQDLN RQRIQRYA
166	1516	Α	2019	2	927	CCQREGLGLKAVVQILLSHGRNGLPGEPASS QGLSAASSTPVFHLALQIDSAPDNIDWVEMLF NKNMVTERLQNVMVLEQCFSDSSSLYRFLTY SYLLAFNVWLLLAPVTLCYDWQVGSIPLVETI WDMRNLATIFLAVVMALLSLHCLAAFKRLE HKEVLVGLLFLVFFFPASNLFFRVGFVVAER VLYMPSMGYCILFVHGLSKLCTWLNRCGATT LIVSTVLLLLLFSWKTVKQNEIWLSRESLFRS GVQTLPHNAKVHYNYANFLKDQGRNKEAIY HYRTALNNNKAWDYLCWRFRKTLTDLP
167	1517	A	2025	696	71	AAASAASSLTVTLGRLASACSHSILRPSGPGA ASLWSASRRFNSQSTSYLPGYVPKTSLSSPPW PEVVLPDPVEETRHHAEVVKKVNEMIVTGQY GRLFAVVHFASRQWKVTSEDLILIGNELDLA CGERIRLEKVLLVGADNFTLLGKPLLGKDLV RVEATVIEKTESWPRIIMRFRKRKNFKKKRIV TTPQTVLRINSIEIAPCLL
168	1518	Α	2046	2	366	HLQVAARVFMPLQAVDSAPKPLKGQAQAPQ RLQGAARVFMPLQAQVKAKASKPLQMQIKA PPRLRRAARVLMPLQAQVRAPRLLQVQSQVS KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR
169	1519	A	2049	1	945	QNLEDREVLNGVQTELLTSPRTKDTLSDMTR TVEISGEGGPLGIHVVPFFSSLSGRILGLFIRGI EDNSRSKREGLFHENECIVKINNVDLVDKTFA QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA NLTGTDSPETDASASLQQNKSPRVPRLGGKPS SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ SGDRILEVNGRDVTGRTQEELVAMLRSTKQG ETASLVIARQEGHFLPRELVMFRSQSH
170	1520	A	2050	363	1	PVATHLTKILNSDEHAVVISSAKTLCETVKDF VAKVEKTYDKTLENAVVADAVASKCSVLNE KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV ESSSEESLGESKEQLGDDVTKPSSQKA
171	1521	Α	2055	139	675	PSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG LDQPGCPKTILGGFQYFKCDHQRSVYLYLSSL

SEQ ID NO: of NO: of Not of	KE
nucle octide sequence	
cotide sequence	
sequence Sequence 09/496 914 09/496 19 19 19 19 19 19 19	
uence 914 ng to first amino acid residue of peptide residue of peptide sequence peptide sequence 914 ng to first amino acid residue of peptide residue of peptide sequence 915 nc peptide sequence 915	
amino acid residue of peptide sequence peptide sequence peptide sequence se	
residue of peptide sequence	
Peptide sequence /=possible nucleotide deletion, \=possible nucleotide insertion RESCTITAYPCDSYQDYRNGKCVSCGTSC SCPLLGYYADNWKDHLRGKDPPMTKAF AEESPFCMYHYFVDIITWNKNVR AEESPFCMYHYFVDIITWNKNVR LIQHKSAVEYAQSHLSLVSMCKESHKCSI MEWKVKIRSDGTRYITKRPVRDRILKER/REERSGLTTDDDTMSEMKMGRY WSKEE HLVRGKEQRRREFFMRIRLKCLKES A 2060 387 GTRILSMQIPFVGFQPIRTSEHMAAAGVF/QAYAFLQYLRDRLTKQEFQTLFFLGVSL/AAVFLSVIYLTYTGYIAPWSGRFYSLWDTCKHIPILASVSEHQPTTWVSFFFDLHILGCTGG G G A 2071 74 443 LLMGPKAKKSGSKKKKVTKAERLKLLQE HRLEAKDLERRNEELEELYLLERCFPEAE HRLEAKDLERRNEELEELYLLERCFPEAE HRLEAKDLERRNEELEELYLLERCFPEAE GETKLLSQWKHYIQCDGSPDPSVAQEMN AALTWSQPQEFWPMEMQPIVTDMVTVH AESSTVGWLCALFRVTHVGVGATGHGVVRRVLCGLPLPSPAPMPIMSLPEGESRKERI RLQFPYLEPGHELPATTLLAFLAAV RRVLCGLPLPSPAPMPIMSLPEGESRKERI RLQFPYLEPGHELPATTLLAFLAAV RRVLCGLPLPSPAPMPIMSLPEGESRKERI RLQFPYLEPGHELPATTLLAFLAAV GIHSVYTVYQGHEIMFNSTMLPYSKENI VERKRHIGNDIVTIVFQEGEESSPAFKPSM FQKFLNLLGDTITLKGWTGYRGGLDTKN GIHSVYTVYQGHEIMFNSTMLPYSKENI VERKRHIGNDIVTIVFQEGEESSPAFKPSM VERKRHIGNDIVTIVFQEGEESSPAFKPSM FQKFLNLLGDTITLKGWTGYRGGLDTKN GIHSVYTVYQGHEIMFNSTMLPYSKENI VERKRHIGNDIVTIVFQEGEESSPAFKPSM FQKFLNLLGDTITLKGWTGYRGGLDTKN FQKFLNLLGDTTTLKGWTGYRGGLDTKN FQKFLNLLGDTTTLKGWTGYRGGLDTKN FQKFLNLLGDTTTLKGWTGYRGGLDTKN FQKFLNLLGDTTTLKGWTGYRGGLDTKN FQKF	
Sequence nucleotide insertion RESCTITAYPCDSYQDYRNGKCVSCGTSQ SCPLLGYYADNWKDHLRGKDPPMTKAF AEESPFCMYHYFVDIIT WNKNVR 172	
RESCTITAYPCDSYQDYRNGKCVSCGTSG SCPLLGYYADNWKDHLRGKDPPMTKAF AEESPFCMYHYFVDIITWNKNVR 172	
SCPLLGYYADNWKDHLRGKDPPMTKAFI AEESPFCMYHYFVDIITWNKNVR 172	
AEESPFCMYHYFVDIITWNKNVR	וטי
172 1522 A 2056 3 361	
MEWKVKIRSDGTRYITKRPVRDRILKERA KEERSGLTTDDDTMSEMKMGRYWSKEE HLVRGKEQRRRREFMMRIRLKCLKES 173 1523 A 2060 1 387 GTRILSMQIPFVGFQPIRTSEHMAAAGVF, QAYAFLQYLRDRLTKQEFQTLFFLGVSLA AVFLSVIYLTYTGYIAPWSGRFYSLWDTG KIHIPIIASVSEHQPTTWVSFFFDLHILGCT G 174 1524 A 2071 74 443 LLMGPKAKKSGSKKKKVTKAERLKLLQE RRLKEEEEARLKYEKEEMERLEIQRIEKEI HRLEAKDLERRNEELEELYLLERCFPEAE QETKLLSQWKHYIQCDGSPDPSVAQEMN AALTWSQPQEFWPMEMQPIVTDMVTVH AESSTVGWLCALFRVTHVGVGATGHGVV RRVLCGLPLPSPAPMPIMSLPEGESRKERE RLQFPYLEPGHELPATTLLAFLAAV 176 1526 A 2092 3 587 EGSVNFKFGVLFAKDGQLTDDEMFSNEIG FQKFLNLLGDTITLKGWTGYRGGLDTKN GIHSVYTVYQGHEIMFHVSTMLPYSKENE VERKRHIGNDIVTIVFQEGEESSPAFKPSM HFTHIFALVRYNQQNDNYRLKIFSEESVPI PPLPTPPVFTDHQEFRDFLLVKLINGEKAT	חע
KEERSGLTTDDDTMSEMKMGRYWSKEE HLVRGKEQRRREFMMRIRLKCLKES	
HLVRGKEQRRREFMMRIRLKCLKES 173 1523 A 2060 1 387 GTRILSMQIPFVGFQPIRTSEHMAAAGVFA QAYAFLQYLRDRLTKQEFQTLFFLGVSLA AVFLSVIYLTYTGYIAPWSGRFYSLWDTG KIHIPIIASVSEHQPTTWVSFFFDLHILGCT G G LMGPKAKKSGSKKKKVTKAERLKLLQE RRLKEEEEARLKYEKEEMERLEIQRIEKE HRLEAKDLERRNEELEELYLLERCFPEAE QETKLLSQWKHYIQCDGSPDPSVAQEMN 175 1525 A 2083 139 486 AALTWSQPQEFWPMEMQPIVTDMVTVH AESSTVGWLCALFRVTHVGVGATGHGV RRVLCGLPLPSPAPMPIMSLPEGESRKERE RLQFPYLEPGHELPATTLLAFLAAV 176 1526 A 2092 3 587 EGSVNFKFGVLFAKDGQLTDDEMFSNEIG FQKFLNLLGDTITLKGWTGYRGGLDTKN GIHSVYTVYQGHEIMFHVSTMLPYSKENE VERKRHIGNDIVTIVFQEGEESSPAFKPSM HFTHIFALVRYNQQNDNYRLKIFSEESVPJ PPLPTPPVFTDHQEFRDFLLVKLINGEKAT	
173 1523 A 2060 1 387 GTRILSMQIPFVGFQPIRTSEHMAAAGVFA QAYAFLQYLRDRLTKQEFQTLFFLGVSLA AVFLSVIYLTYTGYIAPWSGRFYSLWDTG KIHIPIIASVSEHQPTTWVSFFFDLHILGCT G 174 1524 A 2071 74 443 LLMGPKAKKSGSKKKKVTKAERLKLLQE RRLKEEEEARLKYEKEEMERLEIQRIEKEE HRLEAKDLERRNEELEELYLLERCFPEAE QETKLLSQWKHYIQCDGSPDPSVAQEMN 175 1525 A 2083 139 486 AALTWSQPQEFWPMEMQPIVTDMVTVH AESSTVGWLCALFRVTHVGVGATGHGV RRVLCGLPLPSPAPMPIMSLPEGESRKERE RLQFPYLEPGHELPATTLLAFLAAV 176 1526 A 2092 3 587 EGSVNFKFGVLFAKDGQLTDDEMFSNEIG FQKFLNLLGDTITLKGWTGYRGGLDTKN GIHSVYTVYQGHEIMFHVSTMLPYSKENF VERKRHIGNDIVTIVFQEGEESSPAFKPSM HFTHIFALVRYNQQNDNYRLKIFSEESVPJPPLPTPPVFTDHQEFRDFLLVKLINGEKAT	acq
QAYAFLQYLRDRLTKQEFQTLFFLGVSLA AVFLSVIYLTYTGYIAPWSGRFYSLWDTC KIHIPIIASVSEHQPTTWVSFFFDLHILGCT G 174 1524 A 2071 74 443 LLMGPKAKKSGSKKKKVTKAERLKLLQE RRLKEEEEARLKYEKEEMERLEIQRIEKEI HRLEAKDLERRNEELEELYLLERCFPEAE QETKLLSQWKHYIQCDGSPDPSVAQEMN AALTWSQPQEFWPMEMQPIVTDMVTVH AESSTVGWLCALFRVTHVGVGATGHGV RRVLCGLPLPSPAPMPIMSLPEGESRKERE RLQFPYLEPGHELPATTLLAFLAAV 176 1526 A 2092 3 587 EGSVNFKFGVLFAKDGQLTDDEMFSNEIG FQKFLNLLGDTITLKGWTGYRGGLDTKN GIHSVYTVYQGHEIMFHVSTMLPYSKENI VERKRHIGNDIVTIVFQEGEESSPAFKPSM HFTHIFALVRYNQQNDNYRLKIFSEESVPI PPLPTPPVFTDHQEFRDFLLVKLINGEKAT	
AVFLSVIYLTYTGYIAPWSGRFYSLWDTC KIHIPIIASVSEHQPTTWVSFFFDLHILGCT G 174 1524 A 2071 74 443 LLMGPKAKKSGSKKKKVTKAERLKLLQE RRLKEEEEARLKYEKEEMERLEIQRIEKEI HRLEAKDLERRNEELEELYLLERCFPEAE QETKLLSQWKHYIQCDGSPDPSVAQEMN AALTWSQPQEFWPMEMQPIVTDMVTVH AESSTVGWLCALFRVTHVGVGATGHGV RRVLCGLPLPSPAPMPIMSLPEGESRKERE RLQFPYLEPGHELPATTLLAFLAAV 176 1526 A 2092 3 587 EGSVNFKFGVLFAKDGQLTDDEMFSNEIG FQKFLNLLGDTITLKGWTGYRGGLDTKN GIHSVYTVYQGHEIMFHVSTMLPYSKENI VERKRHIGNDIVTIVFQEGEESSPAFKPSM HFTHIFALVRYNQQNDNYRLKIFSEESVPI PPLPTPPVFTDHQEFRDFLLVKLINGEKAT	
KIHIPIIASVSEHQPTTWVSFFFDLHILGCT G 174	
174 1524 A 2071 74 443 LLMGPKAKKSGSKKKKVTKAERLKLLQE RRLKEEEEARLKYEKEEMERLEIQRIEKEI HRLEAKDLERRNEELEELYLLERCFPEAE QETKLLSQWKHYIQCDGSPDPSVAQEMN 175 1525 A 2083 139 486 AALTWSQPQEFWPMEMQPIVTDMVTVH AESSTVGWLCALFRVTHVGVGATGHGVV RRVLCGLPLPSPAPMPIMSLPEGESRKERE RLQFPYLEPGHELPATTLLAFLAAV 176 1526 A 2092 3 587 EGSVNFKFGVLFAKDGQLTDDEMFSNEIG GHSVYTVYQGHEIMFHVSTMLPYSKENE VERKRHIGNDIVTIVFQEGEESSPAFKPSM HFTHIFALVRYNQQNDNYRLKIFSEESVPI PPLPTPPVFTDHQEFRDFLLVKLINGEKAT	
174 1524 A 2071 74 443 LLMGPKAKKSGSKKKKVTKAERLKLLQE RRLKEEEEARLKYEKEEMERLEIQRIEKEI HRLEAKDLERRNEELEELYLLERCFPEAE QETKLLSQWKHYIQCDGSPDPSVAQEMN 175 1525 A 2083 139 486 AALTWSQPQEFWPMEMQPIVTDMVTVH AESSTVGWLCALFRVTHVGVGATGHGV RRVLCGLPLPSPAPMPIMSLPEGESRKERE RLQFPYLEPGHELPATTLLAFLAAV 176 1526 A 2092 3 587 EGSVNFKFGVLFAKDGQLTDDEMFSNEIG FQKFLNLLGDTITLKGWTGYRGGLDTKN GIHSVYTVYQGHEIMFHVSTMLPYSKENE VERKRHIGNDIVTIVFQEGEESSPAFKPSM HFTHIFALVRYNQQNDNYRLKIFSEESVPI PPLPTPPVFTDHQEFRDFLLVKLINGEKAT	PA
RRLKEEEEARLKYEKEEMERLEIQRIEKE HRLEAKDLERRNEELEELYLLERCFPEAE QETKLLSQWKHYIQCDGSPDPSVAQEMN 175 1525 A 2083 139 486 AALTWSQPQEFWPMEMQPIVTDMVTVH AESSTVGWLCALFRVTHVGVGATGHGVV RRVLCGLPLPSPAPMPIMSLPEGESRKERE RLQFPYLEPGHELPATTLLAFLAAV 176 1526 A 2092 3 587 EGSVNFKFGVLFAKDGQLTDDEMFSNEIG FQKFLNLLGDTITLKGWTGYRGGLDTKN GIHSVYTVYQGHEIMFHVSTMLPYSKENE VERKRHIGNDIVTIVFQEGEESSPAFKPSM HFTHIFALVRYNQQNDNYRLKIFSEESVPI PPLPTPPVFTDHQEFRDFLLVKLINGEKAT	
HRLEAKDLERRNEELEELYLLERCFPEAE QETKLLSQWKHYIQCDGSPDPSVAQEMN 175 1525 A 2083 139 486 AALTWSQPQEFWPMEMQPIVTDMVTVH AESSTVGWLCALFRVTHVGVGATGHGVV RRVLCGLPLPSPAPMPIMSLPEGESRKERE RLQFPYLEPGHELPATTLLAFLAAV 176 1526 A 2092 3 587 EGSVNFKFGVLFAKDGQLTDDEMFSNEIG FQKFLNLLGDTITLKGWTGYRGGLDTKN GIHSVYTVYQGHEIMFHVSTMLPYSKENE VERKRHIGNDIVTIVFQEGEESSPAFKPSM HFTHIFALVRYNQQNDNYRLKIFSEESVPI PPLPTPPVFTDHQEFRDFLLVKLINGEKAT	
QETKLLSQWKHYIQCDGSPDPSVAQEMN 175 1525 A 2083 139 486 AALTWSQPQEFWPMEMQPIVTDMVTVH AESSTVGWLCALFRVTHVGVGATGHGVVRRVLCGLPLPSPAPMPIMSLPEGESRKERE RLQFPYLEPGHELPATTLLAFLAAV 176 1526 A 2092 3 587 EGSVNFKFGVLFAKDGQLTDDEMFSNEIG FQKFLNLLGDTITLKGWTGYRGGLDTKN GIHSVYTVYQGHEIMFHVSTMLPYSKENE VERKRHIGNDIVTIVFQEGEESSPAFKPSM HFTHIFALVRYNQQNDNYRLKIFSEESVPJ PPLPTPPVFTDHQEFRDFLLVKLINGEKAT	
175 1525 A 2083 139 486 AALTWSQPQEFWPMEMQPIVTDMVTVH AESSTVGWLCALFRVTHVGVGATGHGVV RRVLCGLPLPSPAPMPIMSLPEGESRKERE RLQFPYLEPGHELPATTLLAFLAAV 176 1526 A 2092 3 587 EGSVNFKFGVLFAKDGQLTDDEMFSNEIG FQKFLNLLGDTITLKGWTGYRGGLDTKN GIHSVYTVYQGHEIMFHVSTMLPYSKENE VERKRHIGNDIVTIVFQEGEESSPAFKPSM HFTHIFALVRYNQQNDNYRLKIFSEESVPI PPLPTPPVFTDHQEFRDFLLVKLINGEKAT	
AESSTVGWLCALFRVTHVGVGATGHGVVRRVLCGLPLPSPAPMPIMSLPEGESRKERE RLQFPYLEPGHELPATTLLAFLAAV 176 1526 A 2092 3 587 EGSVNFKFGVLFAKDGQLTDDEMFSNEIG FQKFLNLLGDTITLKGWTGYRGGLDTKN GIHSVYTVYQGHEIMFHVSTMLPYSKENI VERKRHIGNDIVTIVFQEGEESSPAFKPSM HFTHIFALVRYNQQNDNYRLKIFSEESVPJPPLPTPPVFTDHQEFRDFLLVKLINGEKAT	
RRVLCGLPLPSPAPMPIMSLPEGESRKERE RLQFPYLEPGHELPATTLLAFLAAV 176 1526 A 2092 3 587 EGSVNFKFGVLFAKDGQLTDDEMFSNEIG FQKFLNLLGDTITLKGWTGYRGGLDTKN GIHSVYTVYQGHEIMFHVSTMLPYSKENE VERKRHIGNDIVTIVFQEGEESSPAFKPSM HFTHIFALVRYNQQNDNYRLKIFSEESVPI PPLPTPPVFTDHQEFRDFLLVKLINGEKAT	
RLQFPYLEPGHELPATTLLAFLAAV 176 1526 A 2092 3 587 EGSVNFKFGVLFAKDGQLTDDEMFSNEIG FQKFLNLLGDTITLKGWTGYRGGLDTKN GIHSVYTVYQGHEIMFHVSTMLPYSKENI VERKRHIGNDIVTIVFQEGEESSPAFKPSM HFTHIFALVRYNQQNDNYRLKIFSEESVPI PPLPTPPVFTDHQEFRDFLLVKLINGEKAT	
176 1526 A 2092 3 587 EGSVNFKFGVLFAKDGQLTDDEMFSNEIG FQKFLNLLGDTITLKGWTGYRGGLDTKN GIHSVYTVYQGHEIMFHVSTMLPYSKENI VERKRHIGNDIVTIVFQEGEESSPAFKPSM HFTHIFALVRYNQQNDNYRLKIFSEESVPI PPLPTPPVFTDHQEFRDFLLVKLINGEKAT	VQ
FQKFLNLLGDTITLK.GWTGYRGGLDTKN GIHSVYTVYQGHEIMFHVSTMLPYSKENI VERKRHIGNDIVTIVFQEGEESSPAFKPSM HFTHIFALVRYNQQNDNYRLKIFSEESVPI PPLPTPPVFTDHQEFRDFLLVKLINGEKAT	
GIHSVYTVYQGHEIMFHVSTMLPYSKENI VERKRHIGNDIVTIVFQEGEESSPAFKPSM HFTHIFALVRYNQQNDNYRLKIFSEESVPI PPLPTPPVFTDHQEFRDFLLVKLINGEKAT	SEP
VERKRHIGNDIVTIVFQEGEESSPAFKPSM HFTHIFALVRYNQQNDNYRLKIFSEESVPI PPLPTPPVFTDHQEFRDFLLVKLINGEKAT	
HFTHIFALVRYNQQNDNYRLKIFSEESVPI PPLPTPPVFTDHQEFRDFLLVKLINGEKAT	.QQ
PPLPTPPVFTDHQEFRDFLLVKLINGEKAT	
PPLPTPPVFTDHQEFRDFLLVKLINGEKAT	FG
	LET
177 1527 A 2103 44 427 GKGQVSLEGRPHRGPLCLGSWWPGSRVP	
CDGAWLAWACWVFGNDFPSPASAACSA	
CSVSTACLCVPLCSGSPLAPFRRTAALQE	
RAVSVPLTLAETVASLWPALQELARCGN	AC
RSDLQ .	
178 1528 A 2104 2 409 ALQSTLGAVWLGLLLNSLWKVAESKDQV	
PSTAASSEGAVVEIFCNHSVSNAYNFFWY	
PGCAPRLLVKGSKPSQQGRYNMTYERFS:	
LILQVREADAAVYYCAVEVPNTDKLIFGT	GT
RLQVFPNIQNPD	
179 1529 A 2111 I 312 PTRSSTRPPSLFVHASAKGGEKEEGDDGH	
MRTESHTGLKKGGNANLVFMLKRNTEPK	
SYHFDLERLRAAHILFEREQEHLAPGGISM	PL
PPPLPLACLG	~=-
180 1530 A 2116 3 366 TSIKRAIETTDVTRSFGWDSSEAWQQHDV	
LCRVMFDALEQKWKQTEQADLINELYQO	
KDYVRSLECGYEGWRIDTYLDIPLVIRPY	188
QAFASVVCTFHLTACVSLHRIHNSTVV	
181 1531 A 2117 2 386 , YGLGAHFGRLFIQAGINENDFYDGAWCAI	
NDLQQWIEVDARRLTRFTGVITQGRNSLV	LS
DWVTSYKVMVSNDSHTWVTGKNGSGDN	
GNSEKEIPVLNELPVPMVARYIRINPQSWF	DN
GSICI	
182 1532 A 2123 I 493 RTKTDVYILNLAVADLLLLFTLPFWAVNA	
GWVLGKIMCKITSALYTLNFVSGMQFLAG	
DRYVAVTKVPSQSGVGKPCWIICFCVWM	ISI
LLSIPQLVFYTVNDNARCIPIFPRYLGTSMI	ISI AAI
IQMLEICIGFVVPFLIMGVCYFITARTLMKI	CISI AAI CAL
NIKIS	CISI AAI CAL
183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIF	CISI AAI CAL MP
RSGNYFVTMFDDYSATLPLLIVVILENIAV	CISI AAI CAL MP
VYGIDKFMEDLKDMLGFAPSRYYYYMW	CISI AAI CAL MP VQ CF

CEC ID	L CEO ID	Mai	1 650	Dundiete d	David at James	Ainc sold requests (A - Ali C-C)
SEQ ID	SEQ ID	Met hod	SEQ ID NO:	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
NO: of	NO: of	noa	in NO:	beginning nucleotide	nucleotide location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
1	uence	1	09/496		to last amino	M=Methionine, N=Asparagine, P=Proline,
scq- uence	acrice	1	914	correspondi ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
dence			1714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
]	j		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	Ì	ļ		peptide	i sequence	/=possible nucleotide deletion, \=possible
ĺ			İ	sequence		nucleotide insertion
	 	 	 	- sequence	 -	SPLMLLSLLIASVVNMGLSPPGYNAWIEDKAS
			ì			EEFLSYPTWGLAVCASLDVFAILPVPVAFIGR
			1			RFSLIDDGAGPFCSAAYTTTGCRTPYL
184	1534	A	2145	3	538	HELTVAAADRGQPPQSSVVPVTVTVLDVND
						NPPVFTRASYRVTVPEDTPVGAELLHVEASD
	}			}		ADPGPHGLVRFTVSSGDPSGLFELDESSGTLR
1	})	j	}		LAHALDCETQARHQLVVQAADPAGAHFALA
	}				ĺ	PVTIEVQDVNDHGPAFPLNLLSTSVAENQPPG
	ļ	ļ				TLVTTLHAIDGDAGAFGRLRYHL
185	1535	A	2151	2	671	LDKLLDRMENYNIFNEYILKQVAATYIKLGW
1				_		PKNNFNGSLVQASYQHEELRREVIMLACSFG
1		l	ł	! !		NKHCHQQASTLISDWISSNRNRIPLNVRDIVY
				!		CTGVSLLDEDVWEFIWMKFHSTTAVSEKKIL
1		-				LEALTCSDDRNLLNRLLNLSLNSEVVLDQDAI
	}		1	1		DVIIHVARNPHGRDLAWKFFRDKWKILNTRI
		l				ROKTLEFDFAEPLILAFPILYTAIDNPPLVREH
Ì						E
186	1536	A	2153	2	400	GPMCDKHSAFAEKFHAGFIDYIVHPLWETWA
100	1330	'`	2133	2	100	HLALPDAQDILYTLEDNRNWVDSMIPQSPSPP
İ		1			٠.	LDEQNRDWQGLLENLHVELTLDEEDSEGPEK
1		ļ	ļ			EGEGQTYFTSSKTLCGIVPQNTDSLGETGIHIC
]				AHDKSP
187	1537	A	2158	227	442	FNCFRVASDSFLENSSLLIMILPLRNATQEFIIR
10.	133.	l **	2130	'	174	PGAVAYTCNPSTLGGWGGWITRSGVRDQPG
1			ł			OHGGTPS
188	1538	A	2167	3	486	AHLGGAWLTQRSLGSWAAPGPARAAKEVVA
100	1555	1	2107] ~	100	CIPQNQKMNIWRMKTSKHLQLLSFVLGAVSP
İ		1				AVVVPYMMVLQENGYGVEEGIPTLLMAASS
						MDDILAITGFNTCLSIVFSSGCARSSGSRNSKS
1		1	}			LRTPLGTICEGCDDSSIFSHLDHSSKWSSTYG
İ						HSGA
189	1539	A	2168	2	412	EFLSSNQITQLPNTTFRPMPNLRSVDLSYNKL
105	1555		12,00	7	'''	OALAPDLFHGLRKLTTLHMRANAIOFVPVRIF
İ		į				QDCRSLKFLDIGYNQLKSLARNSFAGLFKLTE
						LHLEHNDLVKVNFAHFPRLISLHSLCLRRNKV
		}	}			AIVVSSLDW
190	1540	A	2179	64	399	MRLNQNTLLLESFGXXRPYTSEHAPTYHQW
		i	1 - 1,5	• •		MKADELLRWTTSEPLTLEHEYAMQRTWLED
						AYECTFIVLDAEKRHAQPGATEESCMVGDVN
1		ļ				LFLTDLEDLTLGEIEVLIAEP
191	1541	A	2190	1	469	CLDRAAGIRHERNVIYINETHTRHRGWLARR
		'`	-1,00	١.	""	LSYVLFIQERDVHKGMFATNVTENVLNSSRV
		[OEAIAEVAAELNPDGSAQQQSKAVNKVKKK
		l				AKRILQEMVATVSPAMIRLTGWVLLKLFNSF
,						FWNIQIHKGQLEMVKAATETNLPLLFLPVHR
İ		[1			SH
192	1542	A	2197	26	157	PSKXGGIRLLLTGTQLYGRFGSAIAPLGDLDR
.,_	15.5	'`	1 217		131	DGYNGEGREEPY
193	1543	A	2236	2	383	EYFPNSIWRSLFSTMDLGDIGFYTYRILQALS
.,,,		1 **	2230	-	203	YTHSKGIMHRDVKPLNILCNSPRNKVILADW
		l			,	GLAEFYHPMRKYSVHVATRYYKSPEILLDYE
						YYDYSLDIWAVGVILLELLTLKLHVFEGGDN
		ł	1			
194	1544	A	2241	105	409	EQ RKGVGKMPTSEGRPGQERSDWVTSYKVMGS
174	1544	^	2241	103	407	NDSHTWVTVKNGSGDMIFEGNSEKEIPVLNE
	l	i			1	LPVPMGARYIRINPOSWFDNGSICMRMEILGC
		!				PLPDPNNY
195	1545	A	2245	1	672	MGVASDWTKRIEYQPGSGSMPLFPSIHLETCD
127	1343	Ι ^ .	4243	1	0/2	GAVSSLQIVTELQTNYIGKGCDRETYSEKSLQ
L		L	L			OWASSTALA LEPATA HOMOCOKEL ISEVOLA

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion KLCGASSGIIDLLPSPSAATNWTAGLLVDSSE MIFKFDGRQGAKIPDGIVPKNLTDQFTITMW MKHGPSPGVRAEKETILCYSDKTEMNRHHY ALYVHNCRLVFLLRKDFDQADTFRPAEFHW KLDQQALAKVDGQPGKSITRQLQEMPVTIQG
196	1546	A	2256	I	396	ISLKPS FRGTPVSGLTNRDTLAVIRHFREPIRLKTVKP GKVINKDLRHYLSLQFQKGSIDHKLQQVIRD NLYLRTIPCTTRAPRDGEVPGVDYNFISVEQF KALEESGALLESGTYDGNFYGTPKPPAEPSPF OPDPV
197	1547	A	2259	43	594	QLAIEIGVRALLFGVFVFTEFLDPFQRVIQPEEI WLYKNPLGQSDNIPTRLMFAISFLTPLAVICV VKIIRRTDKTEIKEAFLAVSLALALNGVCTNTI KLIVGRPRPDFFYRCFPDGVMNSEMHCTGDP DLVSEGRKSFPSIHSSFAFSGLGFTTFYLAGKL HCFTESGRGKSWRLCAAILPL
198	1548	A	2275	3	404	TCTTVVVIPRMLVDFLSESKTISLPECATQMFF FLGFASNNCFIMAAMSYDRYTAIHNPLQYHT LMTRKICLQMMMASWMVGFLFSLCIIVTVFN LSLCDLNTIQHYFCDISPVVSLACNYTFYHEM AIFVLSA
199	1549	A	2315	1	375	LTQMFFIHALSAIESTILLAMAFDRYVAICHPL RHAAVLNNTVTAQIGIVAVVRGSLFFFPLPLLI KRLAFCHSNVLSHSYCVHQDVMKLAYADTL PNVVYGLTAILLVMGXDRMFISLSYFLII
200	1550	Α	2334	2	409	PRVRPQQRKMSFFFKTELGEKLVTKFLFETDF SDDPMLPSPDQLKKKAPFTNKKLKAHQTPVD ILKQKAHQLASMQVQAYNGGNANPRPANNE EEEDEEDEYDYDYESLSDDNILEDRPENKSCH
201	1551	A	2350	3	512	DQLQFEYKEEM ISWEAQIAEIIQWVSDEKDARGYLQALASKM TEELEALRSSSLGSRTLDPLWKVRRSQKLDM SARLELQSALEAEIRAKQLVQEELRKVKDAN LTLESKLKDSEAKNRELLEEMEILKKKMEEK FRADTGKLMLCDSALFEYKYFSNECFYFLFD LIVTLEAPTEFQIQY
202	1552	Α	2351	1	1003	PSSYSSDELSPGEPLTSPPWAPLGAPERPEHLL NRVLERLAGGATRDSAASDILLDDIVLTHSLF LPTEKFLQELHQYFVRAGGMEGPEGLGRKQA. CLAMLLHFLDTYQGLLQEEEGAGHIIKDLYL LIMKDESLYQGLREDTLRLHQLVETVELKIPE ENQPPSKQVKPLFRHFRRIDSCLQTRVAFRGS DEIFCRVYMPDHSYVTIRSRLSASVQDILGSV TEKLQYSEEPAGREDSLILVAVSSSGEKVLLQ PTEDCVFTALGINSHLFACTRDSYEALVPLPE EIQVSPGDTEIHRVEPEDVANHLTAFHWELFR CVHELEFVDYVFHGE
203	1553	A	2361	2	403	NNLNCAEPLFEQNNSLNVNFNTQKKTVWLIH GYRPVGSIPLWLQNFVRILLNEEDMNVIVVD WSRGATTFIYNRAVKNTRKVAVSLSVHIKNL LKHGASLDNFHFIGGSLGAHISGFVGKIFHGQ LGRITGLDP
204	1554	A	2390	280	476	SPSLLPQCLMSLSDLSLSPAPPSHLSPRCPSPQ AGSRLGAMRRCAREMDATPMPPAPSCPSERV T
205	1555	A	2400	543	745	AAVALRDISWQQPYPMDFYAGSSLGPWTVN HGQDRRPHAPGRPARGKVQEGSARPPSAVAC EDCSCR

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion DLSPDSREDHPQGHRRLLPKRPVRGSLMPGH
						THHPCPVSSTTNDTPDQIWVSVGSLRMGTGG MGANASTSPRCWDLSSGNKKWIIQVPILASIV ESRGGLLATGVGGMCACVPRNQPLTGT
207	1557	A	2409	289	418	LWTLYRHKQQVQHNHSNRLSCRPSQEDRAT HTIMVLDKENTLS
208	1558	A	2413	64	492	VQGTGXXFIAFTEAMTHFPASPVWAGMFFL MLINLGLGSMIGTMAGITTPIIDTFKVPKEMFT GGCCVFAFLVGLLFVQRSGNYFVTMFDDYSA TLPLTLIVILENIAVAWIYGTKKFMQELTEML GFRPYRFYFYMWKFVSP
209	1559	A	2417	3	877	EKERLLDEWFTLDEVPKGKLHLRLEWLTLMP NASNLDKVLTDIKADKDQANDGLSSALLILY LDSARNLPIRYKTNEPVWEENFTFFIHNPKRQ DLEVEVRDEQHQCPLGNLKVPLSQLLTSEDM TVSQRFQLGNSGPNSTIKMKIALRVLHLEKRE RPPDHQHSAQVKRPSVSKEGRKTSIKSHMSG SPGPGGSNTAPSTPVIGGSDKPGMEEKAQPPE AGPQGLHDLGRSSSSLLASPGHISVKEPTPSIA SDISLPIATQELRQRLRQLENGTTLGQSPLGQI QLTIP
210	1560	A	2422	35	456	REFAASDLEPFTPTDQPISPEAITQPSCIKRQRA AGNPGSLAATIDHKPCSAPLEPKIQASRNQRW GAVRAAESLTDIAEPASPQVHETPIDASQTQK VEPASKSRFTPELQAKVSHSRERALSTMDATP HHAQPQRGEG
211	1561	Α	2431	1	764	RRYSOKLIOHTACQLLRTYPAATRIDSSNPNP LMFWLHGIQLVALNYQTDDLPLHLNAAMFE ANGGCGYVLKPPVLWDKNCPMYQKFSPLER DLDSMDPAVYSLTIVSGQNVCPSNSMGSPCIE VDVLGMPLDSCHFRTKPIHRNTLNPMWNEQF LFHVHFEDLVFLRFAVVENNSSAVTAQRIIPL KALKRGYRHLQLRNLHNEVLEISSLFINSRRM EENSSGNTMSASSMFNTEERKCLQTHRVTVH GVPG
212	1562	A	2436	1	411	GIRGTTGHLGCPINDDPSLTLTVSWVMEDKPI YIGNGTKKEDDSLTIFAVAKRDHVSDTCGAC TDLDHNLDKGYLTVLGEQATPTNRLGALPKG RANRTRDLELTYLAERIVRLTWIPGDANNRPI TDYDCQIEEHQ
213	1563	A	2445	1		MSSIGCL WVSRSSQIDGLTAEKSGPEKPHGT WLMPELHPKEQILELLVLEQFLSILPEELQIWV QQHNPESGEESVTLLEDLEREFDDPGQQVPAS PQGPAVPWKDLTCLRASQESTDIHLQPLKTQ LKSWKPCLSPKSDCENSETATKEGISEEKSQG LPQEPSFRGISEHESNLVWKQGSATGEKLRSP SQGGSFSQVIFTNKSLGKRDLYDEAERCLILT TDSIMCQK VPPEERPYRCDVCGHSFKQHSSLT QHQRIHTGEKPYKCNQCGKAFSLRSYLIHQR IHSGEKAYECSECGKAFNQSSALIRHRKIHTG EKACKCNECGKAFSQSSYLIHQRIHTGEKPY ECNECGKTFSQSSKLIRHQRIHTGERPYECNE CGKAFRQSSELITHQRIHSGEKPYECSECGKA FSLSSNLIRHQRIHSG
214	1564	A	2461	1	615	GIPGSTISSSRNIFLEDDLAWQSLIHPDSSNTPL STRLVSVQEDAGKSPARNRSASITNLSLDRSG SPMVPSYETSVSPQANRTYVRTETTEDERKIL LDSVQLKDLWKKICHHSSGMEFQDHRYWLR THPNCIVGKELVNWLIRNGHIATRAQAIAIGQ

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl- cotide	peptide		in USSN	nucleotide location	location	F=Phenylalanine, G=Glycine, H=Histidine,
seq-	seq- uence		09/496	correspondi	corresponding to last amino	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
uence	uence	ļ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
			1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			ļ	peptide		/=possible nucleotide deletion, \=possible
	<u> </u>			sequence		nucleotide insertion
			İ	1		AMVDGRWLDCVSHHDQLFRDEYALYRPLQV
215	15.05	<u> </u>	0464	<u> </u>	0000	LFSVYCQLECSKLIL
215	1565	A	2464	3	2932	GPGVRSSQDGMADVFVHLRTAWPRCSFISGQ
)					ļ	HGPGRHGRRVCSSQDSMADVFVHLRTAWPT
				•		CSLISGQHGPGESVSYEDDDIPAPASLLHVNA AAPALTNPTAPVLCTAPNNTAQKEKVPSGMR
				Ì		QRPAGVRISSRTPDLTCAVSTHSTVPGVRISSC
1	}	l				TPDLTCAVSIHSTVPSVCISSCTPDLTCAVSTH
					•	STVPGVRISSCTPDLTCAVSTHSTVPGVRISSR
						TPDLTCAVSIHATVPGVRISSCTPDLTCAVSIH
						ATVPGVRISSCTPDLTCAVSTHSTVPGVRISSR
		1		}		TPDLTCAVSIHSTVPGVRISSCTPDLTCAVSIH
1		1	l .			ATVPGVRISSCTPDLTCAVSTHSTVPGVRISSR
	1					TPDLTCAVSIHATVPGVRISSRTPDLTCAVSIH
						ATVPGVRISSCTPDLTCAVSIHATVPGVRISSC TPDLTCAVSIHATVPGVRISSRTPDLTCAVSIH
		[ATVPGVRISSCTPDLTCAVSTHSTVPGVRISSR
	1	Ì				TPDLTCAVSIHATVPGVRISSCTPDLTCAVSTH
1	1	ļ				STVPGVRISSRTPDLTCAVSIHATVPGVHISSC
	i		!			TPDLTCAVSTHSTVPGVRISSRTPDLTCAVSIH
						STVPGVCISSRTPDLTCAVSIHSTVPSVHISSCT
	1					PDLTCAVSIHSTVPGVRISSRTPDLTCAVSTHS
						TVPGVHISSCTTDLTCAVSIHATVPGVHISSCT
						PDLTCAVSTHTTVPGVRISSRTPDLTCAVSIHS TVPGVRISSCTPDLTCAVSTHSTVPGVRISSRT
		Ì				PDLTCAVSTHLTVPGVRISSRTPDLTCAVSIHA
						TVPGVHISSCIPDLTCAVSIHATVPGVRISSRT
j]]		,	PDLTCAVSIHATVPGVHISSCTPDLTCAVSTHS
	İ					TVPGVRISSRTPDLTCAVSIHSTVPGVHISSCT
]					PDLTCAVSTHSTVPGVHISSCTPDLTCAVSTH
						STVPGVHISSRTPDLTCAVSIHATVPSVHISSC
216	1566	A	2477	•	414	TPDLTCAVSIHSTVPGLLTSVSQTSTG
210	1300	A	24//	1	414	FRTKSYRKGSYRCIVSEWIAEQGNWQEIQEK
		ĺ	[[AVEVATVVIQPTVLRAAVPKNVSVAEGKELD LTCNITTDRADDVRPEVTWSFSRMPDSTLPGS
1]			RVLARLDRDFLVHSSPHVALSHVDARSYHLL
1]]			VRDVSKENSGYYY
217	1567	Α	2480	2	460 ·	CRTLCEGPQRFEEYEYLGYKAGLYEAIADHY
		l			1	MQVLVCQHECVRELATRPGRLSPIENFLPLHY
1					l	DYLQFAYYRVGEYVKALECAKAYLLCHPDD
				1		EDVLDNVDYYESLLDDSIDPASIEAREDLTMF
210	1560		2402	140	202	VKRHKLESELIKSAAEGLGXSYTEPNYW
218	1568	A	2483	140	383	AFSSPHPSPAPQFPECGFYGLYDKILLFKHDPT
				ì		SANLLQLVRSSGDIQEGDLVEVVLSASATFED LQIRPHALTVHSYRAP
219	1569	A	2489	3	428	SSRLVLLAGAAALASGSQGDREPVYRDCVLO
	1507	71	2.07	·	120	CEEQNCSGGALNHFRSRQPIYMSLAGWTCRD
						DCKYECMWVTVGLYLQEGHKVPQFHGKWP
1						FSRFLFFQEPASAVASFLNGLASLVMLCRYRT
L						FVPASSPMYHTCVAFAWVS
220	1570	A	2498	1	1297 .	MDGEAVRFCTDNQCVSLHPQEVDSVAMAPA
1]			ı		APKIPRLVQATPAFMAVTLVFSLVTLFVVDH
1					•	HHFGREAEMRELIQTFKGHMENSSAWVVEIQ
ļ	}					MLKCRVDNVNSQLQVLGDHLGNTNADIQMV
				[KGVLKDATTLSLQTQMLRSSLEGTNAEIQRL
]	İ		ļ		KEDLEKADALTFQTLNFLKSSLENTSIELHVL
1		1		l	l	SRGLENANSEIQMLNASLETANTQAQLANSS LKNANAEIYVLRGHLDSVNDLRTQNQVLRNS
			ļ	1	j	LEGANAEIQGLKENLQNTNALNSQTQAFIKSS
	1					

C 250 15	1 000 10	1377	Long	1.5	1	
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning nucleotide	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	Į	in		location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN 09/496	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	иепсе	ł	(correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1		ĺ		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible nucleotide insertion
		 		sequence		FDNTSAEIQFLRGHLERAGDEIHVLKRDLKM
						VTA OTOV ANCHI DOTESTOIO VEVER VENEZI
						VTAQTQKANGRLDQTDTQIQVFKSEMENVN TLNAQIQVLNGHMKNASREIQTLKQGMKNA
		l		l		SALTSQTQMLDSNLQKASAEIQTLRQGMKNA
221	1571	A	2501	3	500	KALTMEIQQEQSRLKTLHVVITSQEQLQRTQ
221	13/1	l A	2301	3	300	RVRLNNDGLSPLMMAAKTGKIGIFQHIIRREV
		ļ		İ		TDEDTRHLSRKFKDWAYGPVYSSLYDLSSLD
{	[!	TCGEEASVLEILVYNSKIENRHEMLAVEPINE
					İ	LLRDKWRKFGAVSFYINVVSYLCAMVIFTLT
		ļ				AYYQPLEGTPPYPYRTTVDYLRLAGEVITLFT GVLFFFTN
222	1572	 _,	0500	1	206	
222	15/2	Α	2508	3	395	DAHCQRKLAMQEFMEINERLTELHTQKQKL
		ļ		}		ARHVRDKEEEVDLVMQKVESLRQELRRTER
	ł	l	ł			AKKELEVHTEALAAEASKDRKLREQSEHYSK
						QLENELEGLKQKQISYSPGVCSIEHQQEITKL
-002	1500	<u> </u>	0544		410	KTDLEKKS
223	1573	Α	2544	2	412	NDPAIISNESAAV VHTIVNETLESMESLEVTK
ļ		İ		•		MVDERTDYLTKSLKEKTPPFSHCDQAVLQCS
ĺ	1			ĺ		EASSNKDMFADRLSKSIIKHSIDKSKSVIPNID
					KNAVYKESLPVSGEESQLTPEKSPKFPDSQNQ	
004	1001	ļ.,	-	101		LTHCSLSAA
224	1574	Α	2552	401	l	GASLCFISTAFTVLTFLIDSCRFSYPERPIIFLSM
				i		CYNIYSIAYIVRLTVGRERISCDFEEAAEPVLI
ļ ·		ļ				QEGLKNTGCAIIFLLMYFFGMASSIWWVILTL
	ì	1	ł	ì		TWFLAAGLKWGHEAIEMHSSYFHIAAWAIPA
225	1575	<u> </u>	2563	724	1	VK ,
223 .	13/3	Α	2363	124	1	MSARKERREKGEEEGEGEKDGDEDEKEEEKE
			1			GLGEEEKEAGKKKKKQEEKEKGAVYSR
						VARICKNDMGGSQRVLEKHWTSFLKARLNC
		ĺ				SVPGDSFFYFDVLQSITDIIQINGIPTVVGVFTT
						QLNSIPGSAVCAFSMDDIEKVFKGRFKEQKTP DSVWTAVPEDKVPKPRPGCCAKHGLAEAYK
						TSIDFPDETLSFIKSHPLMDSAVPPIADEPWFT
		i				KTRVRYRLTAISVDHSAGPYH
226	1576	A	2571	449	3	EGVLFVYGNYVGDVMNFEMAAEMAQEVAIP
220	1376	Α .	23/1	449	3	TRTVLTTDDISSSPIEDRDGRRGVAGNFFIFKV
		İ				AGAACDRGMSLEACEAVTRKANRRTYTMG
						VALEPCSLPQTRRHNFEIGAEEMEIGMGIHGE
227	1577	A	2575	3	1197	RGVIREKMMPADAIVDHIMDRIFS VLSDLCLFYYRDEKEEGILGSILLPSFQIALLTS
221	1311	Γ.	6167	ا	1177	
		1				EDHINRKYAFKAAHPNMRTYYFCTDTGKEM
		}				ELWMKAMLDAALVQTEPVKRVDKITSENAP
						TKETNNIPNHRVLIKPEIQNNQKNKEMSKIEE KKALEAEKYGFQKDGQDRPLTKINSVKLNSL
		1				PSEYESGSACPAQTVHYRPINLSSSENKIVNVS LADLRGGNRPNTGPLYTEADRVIQRTNSMQQ
						LEQWIKIQKGRGHEEETRGVISYQTLPRNMPS
		[1		!	HRAOIMARYPEGYRTLPRNSKTRPESICSVTP
						STHDKTLGPGAEEKRRSMRDDTMWQLYEW
						QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT
					İ	MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI
						QRGDVTIDRRHRAHHPKVK
228	1578		2583		220	
220	13/8	A	2303	3	330	LPFLGLGSVLPQGMVMASPEMNPTICSVFEA
	·	1				HIVLLFHATTFRRGFQVTVLVGNVRQTAVVE
	.		1 1			KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD
229	1579	A	2589	1	440	PTMGIKPHLWWVAA
227	13/3	A .	2309	1	448	DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR
		ľ	į l			ECVAPNICKCKPGYIGSNCQTALCDPDCKNH
		L	<u> </u>			GKCIKPNICQCLPGHGGATCDEEHCNPPCQH

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid. E=Glutamic Acid.
nucl-	peptide	1100	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine.
eotide	seq-	J	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ĺ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	uence	1	914		acid residue	
uence		1	914	ng to first		Q=Glutamine, R=Arginine, S=Serine,
ł	İ	Į	Ì	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
ĺ		ĺ	Î	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			1	peptide		/=possible nucleotide deletion, \=possible
<u></u>		<u> </u>	<u> </u>	sequence		nucleotide insertion
ŀ		}		1		GGTCLAGNLCTCPYGFVGPRCETMVCNRHC
		<u> </u>		<u> </u>		ENGGQCLTPDICQCKPGWYGPTCSTA
230	1580	A	2593	2	138	AVTFSVVFAYVADITQEHERSMAYGLVCMFI
		<u></u>		L_	L	LYLLYLLRNAFFLR
231	1581	Α	2595	185	2	SGPYTDFTPWPTEEQKLLEQALKTYPVNPPER
İ	4		ľ	ĺ	i	WEKIAEAVPGRTKKACIKRYKVADLRISK
232	1582	Α	2596	1	391	STVTGQPRRLLDTAGHQQPFLELKIRANEPGA
i	i	i	i	İ	:	GRARRATPTCEPATPLCCRRDHYVNFQELGW
ł		i		1		RDWILLPEGYQLNYCSGQCPTHLAGSPGIAAS
ı	ł	}	ł	ł		FHSAVFSLLKANNPWPGRTSWCVPTARRPLS
	!					LLYL
233	1583	A	2601	184	403	LLFSDEIIMAAPLRIADVTSGLIGGEDGRVYV
233	1565	^	2001	104	403	
	1		1			YNGKETTLGDMTGKCKSWITPCPEEKVNVLQ
1004	1.504	<u> </u>		<u></u>	ļ	NSIPYWERIT
234	1584	Α	2614	178	335	PLTLCLPENNKPPQADAVPDKELTLPVDSTTL
		Ļ				DGSKSSDDQKIISYLWEKTQ
235	1585	A	2616	2	896	DVLEVYGTGVASTRHEMGTLDKHKELEDLV
ł						AKFLNVEAAMVFGMGFATNSMNIPALVGKG
ļ			}	1		CLILRDEVNHTSLVLGARLLGATIGIFKHNYA
į		İ	l.			QSLEKLLRDAVIYGQPRTRRAWKKILILVEGV
ŀ	ļ	ļ	,			YSMEGSIVHLPQIIALKKKYKAYLYIDEAHSI
1	1		{			GAVGPTGRGVTEFFGLDPHEVDVLMGTFTKS
1		į	ŀ			FGASGGYIAGRKARILSPPACLVPNTGSHSLH
	{	1		1	(RLTRDLQMNEAMVALVTDRLQGWNSGEGN
	l		ł			WDRADKFGDLVDYLRVHSHSAVYASSMSPPI
			į			AEQIIRSLKLIMGLDGTTQ
236	1586	A	2621	 	392	NTSSFPAQPSSPARPSLPHLSQHPSNPLLPLAS
200	1300	1	2021	1.	372	ADHPQCGRFLPLHEPEPLCPSPSLSYPTLVSS
1	J	j	j	j		WSSPFSSHHGCPPGLYPFPTSPKTIQPPGLAQL
1						KMLCIPPGRQQLRGAQSMPGHGALSPLLLPP
						A
237	1587	A	2628	398	1	DLVCKISGFGRGPRDRSEAVYTTMSGRSPAL
237	1307	A	2020	398	1	
	1			j	1 1	
	1	1		1		WAAPETLQFGHFSSASDVWSFGIIMWEVMAF
		İ				GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN
						GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV
						GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV
238	1588	A	2631	1	1104	GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV
238	1588	A	2631	1	1104	GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV
238	1588	A	2631	1	1104	GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV
238	1588	A	2631	ì	1104	GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ
238	1588	A	2631	1	•	GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW
238	1588	A	2631	1	•	GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS
238	1588	A	2631	1	•	GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI
238	1588	A	2631	1	•	GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA
238	1588	A	2631	1	•	GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT
238	1588	A	2631	1	•	GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF
238	1588	A	2631	1	•	GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF VIKLIGGNRKLVARPLSPRSEEEVLAGRKGGP
238	1588	A	2631	1	:	GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF VIKLIGGNRKLVARPLSPRSEEEVLAGRKGGP KEALQTHKHQNGIFSNGSKAEKRGLAANPGS
					:	GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF VIKLIGGNRKLVARPLSPRSEEEVLAGRKGGP KEALQTHKHQNGIFSNGSKAEKRGLAANPGS RYDDLVSRLLEQGAPCSSSKKKN
238	1588	A	2631	1	:	GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF VIKLIGGNRKLVARPLSPRSEEEVLAGRKGGP KEALQTHKHQNGIFSNGSKAEKRGLAANPGS RYDDLVSRLLEQGAPCSSSKKKN MKPDNILLDEHGHVHITDFNIAAMLPRETQIT
					:	GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT VAPFGYLKHRLKPSDAGVYTCSAGPAREHF VIKLIGGNRKLVARPLSPRSEEEVLAGRKGGP KEALQTHKHQNGIFSNGSKAEKRGLAANPGS RYDDLVSRLLEQGAPCSSSKKKN MKPDNILLDEHGHVHITDFNIAAMLPRETQIT TMAGTKPYMAPEMFSSRKGAGYSFAVDWW
					. :	GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF VIKLIGGNRKLVARPLSPRSEEEVLAGRKGGP KEALQTHKHQNGIFSNGSKAEKRGLAANPGS RYDDLVSRLLEQGAPCSSSKKN MKPDNILLDEHGHVHITDFNIAAMLPRETQIT TMAGTKPYMAPEMFSSRKGAGYSFAVDWW SLGVTAYELLRGRRPYHIRSSTSSKEIVHTFET
					. :	GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT VAPFGYLKHRLKPSDAGVYTCSAGPAREHF VIKLIGGNRKLVARPLSPRSEEEVLAGRKGGP KEALQTHKHQNGIFSNGSKAEKRGLAANPGS RYDDLVSRLLEQGAPCSSSKKKN MKPDNILLDEHGHVHITDFNIAAMLPRETQIT TMAGTKPYMAPEMFSSRKGAGYSFAVDWW
					. :	GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF VIKLIGGNRKLVARPLSPRSEEEVLAGRKGGP KEALQTHKHQNGIFSNGSKAEKRGLAANPGS RYDDLVSRLLEQGAPCSSSKKN MKPDNILLDEHGHVHITDFNIAAMLPRETQIT TMAGTKPYMAPEMFSSRKGAGYSFAVDWW SLGVTAYELLRGRRPYHIRSSTSSKEIVHTFET
					. :	GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF VIKLIGGNRKLVARPLSPRSEEEVLAGRKGGP KEALQTHKHQNGIFSNGSKAEKRGLAANPGS RYDDLVSRLLEQGAPCSSSKKN MKPDNILLDEHGHVHITDFNIAAMLPRETQIT TMAGTKPYMAPEMFSSRKGAGYSFAVDWW SLGVTAYELLRGRRPYHIRSSTSSKEIVHTFET TVVTYPSAWSQEMVSLLKKLLEPNPDQRFSQ
					. :	GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF VIKLIGGNRKLVARPLSPRSEEEVLAGRKGGP KEALQTHKHQNGIFSNGSKAEKRGLAANPGS RYDDLVSRLLEQGAPCSSSKKKN MKPDNILLDEHGHVHITDFNIAAMLPRETQIT TMAGTKPYMAPEMFSSRKGAGYSFAVDWW SLGVTAYELLRGRRPYHIRSSTSSKEIVHTFET TVVTYPSAWSQEMVSLLKKLLEPNPDQRFSQ LSDVQNFPYMNDINWDAVFQKRLIPGFIPNK
					. :	GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF VIKLIGGNRKLVARPLSPRSEEEVLAGRKGGP KEALQTHKHQNGIFSNGSKAEKRGLAANPGS RYDDLVSRLLEQGAPCSSSKKN MKPDNILLDEHGHVHITDFNIAAMLPRETQIT TMAGTKPYMAPEMFSSRKGAGYSFAVDWW SLGVTAYELLRGRRPYHIRSSTSSKEIVHTFET TVVTYPSAWSQEMVSLLKKLLEPNPDQRFSQ LSDVQNFPYMNDINWDAVFQKRLIPGFIPNK GRLNCDPTFELEEMILESKPLHKKKKRLAKK
					. :	GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF VIKLIGGNRKLVARPLSPRSEEEVLAGRKGGP KEALQTHKHQNGIFSNGSKAEKRGLAANPGS RYDDLVSRLLEQGAPCSSSKKN MKPDNILLDEHGHVHITDFNIAAMLPRETQIT TMAGTKPYMAPEMFSSRKGAGYSFAVDWW SLGVTAYELLRGRRPYHIRSSTSSKEIVHTFET TVVTYPSAWSQEMVSLLKKLLEPNPDQRFSQ LSDVQNFPYMNDINWDAVFQKRLIPGFIPNK GRLNCDPTFELEEMILESKPLHKKKKRLAKK EKDMRKCDSSQTCLLQEHLDSVQKEFIINRE KVNRDCI
239	1589	A	2636	1	678	GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF VIKLIGGNRKLVARPLSPRSEEEVLAGRKGGP KEALQTHKHQNGIFSNGSKAEKRGLAANPGS RYDDLVSRLLEQGAPCSSSKKN MKPDNILLDEHGHVHITDFNIAAMLPRETQIT TMAGTKPYMAPEMFSSRKGAGYSFAVDWW SLGVTAYELLRGRRPYHIRSSTSSKEIVHTFET TVVTYPSAWSQEMVSLLKKLLEPNPDQRFSQ LSDVQNFPYMNDINWDAVFQKRLIPGFIPNK GRLNCDPTFELEEMILESKPLHKKKKRLAKK EKDMRKCDSSQTCLLQEHLDSVQKEFIIINRE KVNRDCI ELLDPTTPMRTKCIELLYAALTSSSTDQPKAD
239	1589	A	2636	1	678	GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF VIKLIGGNRKLVARPLSPRSEEEVLAGRKGGP KEALQTHKHQNGIFSNGSKAEKRGLAANPGS RYDDLVSRLLEQGAPCSSSKKN MKPDNILLDEHGHVHITDFNIAAMLPRETQIT TMAGTKPYMAPEMFSSRKGAGYSFAVDWW SLGVTAYELLRGRRPYHIRSSTSSKEIVHTFET TVVTYPSAWSQEMVSLLKKLLEPNPDQRFSQ LSDVQNFPYMNDINWDAVFQKRLIPGFIPNK GRLNCDPTFELEEMILESKPLHKKKKRLAKK EKDMRKCDSSQTCLLQEHLDSVQKEFIINRE KVNRDCI

241	SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion EMANKELKQLRASYTESCIQEHYLPQVIDGTL Y
MYLMAGAKPOMILLCFMCCTTILS STTAMVMPIVEAVLQELVSAEDEQLY	241	1591	A		392	3	IRLTILRCVFMRLATICVLVFTLGSKITSCDDD TCDLCGYNQKLYPCWETQVGQEMYKLMIFD FIIILAVTLFVDFPRKLLVTYCSSCKLIQCWGQ QEFAIPDNVLGIVYGQTICWIGAFFSPLLPAM
LINGSIDGIRHMFTPKLEIMLEPKVW VFFALGLGFGGV1AFSSYNRDNNCH VSFINFTISVLATIVVFAVLGFKANVI VSFRIFTSVLATIVVFAVLGFKANVI VSFRIFTSVLATIVVFAVLGFKANVI VSFRIFTSVLATIVVFAVLGFKANVI QNSETV							
AAVLMLLMCIFALIAHWLACIWYAIG YI.TDKIGWLDSLQQIGKRYNDSDS DKYVTALYFIFSSLTSVGFGNVSPNT SICVMLIGSLMYASIFGNVSAIIQRLYS HMQMLRVKEFIRFHQIJPNPLRQRLES WTYTNIGDMINVTNIGTCSSCTSDIG NHHQGGLIYSWNDAASMQRPFNHIK; TSDSNLNKYSTINKIPQLILNFSEVKTI PPSSDKTIIAPKVKDRTHINVTEKVTQ DVLPEYKLQAPRINKFTILHYSPFKAV LLLVIYTAIFTPYSAAFLLNDREEQKR SCSPLNVVDLIJMFIIDILINFRITTY VVSDPASV 245 1595 A 2656 385 2 NLTWWPLFRDVSFYIVDLIMLIIFFLD WESLILLTAYFCYVVFMKFNVQVEK MIRNKVVKVTAPEAQAKPSAARDK AKPRLQRGGSSASLHNSLMRNSIFQN PHV							CLAMIKGIQSSGKIIYFSSLFPYVVLICFLIRAF LLNGSIDGIRHMFTPKLEIMLEPKVWREAATQ VFFALGLGFGGVIAFSSYNKRDNNCHFDAVL VSFINFFTSVLATLVVFAVLGFKANVINEKCIT QNSETV
WESLLLTAYFCYVVFMKFNVQVEK MINRNKVVKVTAPEAQAKPSAARDK AKPRLQRGGSSASLHNSLMRNSIFQNI PHV	244	1594	A	2650	1	1271	MTTTLIGLLKTARLLRLVRVARKLDRYSEYG AAVLMLIMCIFALIAHWLACIWYAIGNVERP YLTDKIGWLDSLGQQIGKRYNDSDSSSGPSIK DKYVTALYFTFSSLTSVGFGNVSPNTNSEKIF SICVMLIGSLMYASIFGNVSAIIQRLYSGTARY HMQMLRVKEFIRFHQIPNPLRQRLEEYFQHA WTYTNGIDMNMVTNGTCSSCTSDDGHFILVS NHHQGGLIYSWNDAASMQRPFNHIKSSLLGS TSDSNLNKYSTINKIPQLTLNFSEVKTEKKNSS PPSSDKTIIAPKVKDRTHNVTEKVTQVLSLGA DVLPEYKLQAPRINKFTILHYSPFKAVWDWLI LLLVIYTAIFTPYSAAFLLNDREEQKRRECGY SCSPLNVVDLIVDIMFIIDILINFRTTYVNQNEE VVSDPASV
LLDMRIKHLIKTNQLSQATALAKLCSI IKGSFKQTYLVCLCTSSPNGKLIEEVSN NYFLS 247 1597 A 2678 3 267 DAWVKNDIIFNQTERKQKISENLKHL/ VQKNLVFVVGLSQRLADPEVSPLVFF VSLSYLEIIFDPAQLCDSSEHIIS 248 1598 A 2687 1 404 DFTTLAAMMRTLFSLFGDVRSDVHRF GAAIKSVKNPDKKSIENQVLDSLVPLL ENDAVAEESRQVLTICAQFLKWKLPR DPWHIKPTEAGTICRFFEKKCKGKINII MYSKNPKL 249 1599 A 2692 1 440 FRRRRRRERDCAAQGARRHCRHLAI SFPIGIYKVLRNVSGQIHLITLANNELK FMTTFSQLRELHLEGNFLHRLPSEVSA AIDLSRNQFQDFPEQLTALPALETINLE DVPVEKLAAMPALRSINL 250 1600 A 2693 459 21 LLPGSLGVPILHSQPWDPSPQCPHRAP PPLGALSQALTFLSRAAKNHSQDPGKO	245	1595	A	2656	385	2	NLTWWPLFRDVSFYIVDLIMLIIFFLDNVIMW WESLLLLTAYFCYVVFMKFNVQVEKWVKQ MINRNKVVKVTAPEAQAKPSAARDKDEPTLP AKPRLQRGGSSASLHNSLMRNSIFQNKIHTLD PHV
VQKNLVFVVGLSQRLADPEVSPLVFF VSLSYLEIIFDPAQLCDSSEHIIS 248 1598 A 2687 1 404 DFTTLAAMMRTLFSLFGDVRSDVHRF GAAIKSVKNPDKKSIENQVLDSLVPLL ENDAVAEESRQVLTICAQFLKWKLPR DPWHIKPTEAGTICRFFEKKCKGKINII MYSKNPKL 249 1599 A 2692 1 440 FRRRRRERDCAAQGARRHCRHLAI SFPIGIYKVLRNVSGQIHLITLANNELK FMTTFSQLRELHLEGNFLHRLPSEVSA AIDLSRNQFQDFPEQLTALPALETINLE DVPVEKLAAMPALRSINL 250 1600 A 2693 459 21 LLPGSLGVPILHSQPWDPSPQCPHRAP PPLGALSQALTFLSRAAKNHSQDPGKO	246	1596	A	2660	200	506	VLVLQMNYYQMLIIYYVLFFKVNEFLAFEGPI LLDMRIKHLIKTNQLSQATALAKLCSDHPEIG IKGSFKQTYLVCLCTSSPNGKLIEEVSMFSFIS NYFLS
GAAIKSVKNPDKKSIENQVLDSLVPLI ENDAVAEESRQVLTICAQFLKWKLPR DPWHIKPTEAGTICRFFEKKCKGKINII MYSKNPKL 249 1599 A 2692 1 440 FRRRRRRERDCAAQGARRHCRHLAI SFPIGIYKVLRNVSGQIHLITLANNELK FMTTFSQLRELHLEGNFLHRLPSEVSA AIDLSRNQFQDFPEQLTALPALETINLE DVPVEKLAAMPALRSINL 250 1600 A 2693 459 21 LLPGSLGVPILHSQPWDPSPQCPHRAP PPLGALSQALTFLSRAAKNHSQDPGKO	247		Α	2678	3	267	DAWVKNDIIFNQTERKQKISENLKHLASVRV VQKNLVFVVGLSQRLADPEVSPLVFFVILIFF VSLSYLEIIFDPAQLCDSSEHIIS
SFPIGIYKVLRNVSGQIHLITLANNELK FMITTFSQLRELHLEGNFLHRLPSEVSA AIDLSRNQFQDFPEQLTALPALETINLE DVPVEKLAAMPALRSINL 250 1600 A 2693 459 21 LLPGSLGVPILHSQPWDPSPQCPHRAP PPLGALSQALTFLSRAAKNHSQDPGKO			Ā	2687	1	404	DFTTLAAMMRTLFSLFGDVRSDVHRFSVTLF GAAIKSVKNPDKKSIENQVLDSLVPLLLYSQD ENDAVAEESRQVLTICAQFLKWKLPREVYSK DPWHIKPTEAGTICRFFEKKCKGKINILEQTL MYSKNPKL
PPLGALSQALTFLSRAAKNHSQDPGKO		,	Ā	2692	1	440	
NSPWHPATLPGALGPQLSQAAPSPIPPI ISSCPDLKLTKSSTP							LLPGSLGVPILHSQPWDPSPQCPHRAPSTPRRL PPLGALSQALTFLSRAAKNHSQDPGKGTKPFP AAPAAPPRSSLPAPLPMGLKDKGPQPAPPTIF NSPWHPATLPGALGPQLSQAAPSPIPPPCLMG ISSCPDLKLTKSSTP FVFDLKLRVPGFAALLIHGASSVPGPETVRLR

Deptide	SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine,
Double			1100	1			D=Aspartic Acid, E=Glutamic Acid,
Sequence		1		1			
1914 maino acid rasidue of peptide peptide peptide peptide peptide peptide peptide peptide sequence			}				M=Methionine, N=Asparagine P=Proline
mino acid residue of peptide residue of peptide sequence T-Threonine, V-Valine, W-Toptophen, y-Expression condon, peptide sequence T-Threonine, V-Valine, W-Toptophen, y-Tyrosia, X-Unknown, **Sepo codon, p-possible nucleotide deletion, \(\text{possible} \) Possible nucleotide insertion QKRKKKAPJHSSGRKEELVTTHT/DKLETKK PKRKKAGFEELVTHT/DKLETKKAGFEELVTHT/DKLETKKAGFEELVTHT/DKLETKKAGFEELVTH/DKLETKAGFEELVTH/DKLETKKAGFEELVTH/DKLETKKAGFEELVTH/DKLETKAGFEELVTH/DKLE	uence			914		1	Q=Glutamine, R=Arginine, S=Serine.
Peptide Sequence	ĺ		l		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan.
	1)	ļ		sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	1		1				/=possible nucleotide deletion, \=possible
PYGRVLCGLSGELHSLIJPRKYTEKRALGSH RAGFFEHPVAPPELSNSCOJSKEGREQVISE GAGDCL		ļ		<u> </u>	sequence		
RRAGFPEHPVAPERISNSCOJSKEGREQVLSE GADDC G	}	Ì					QKRKKKAPDHSSGRKEELVTTHTVDKLETKK
GAGDCL		ł	l	ļ		1	
252							
1603	252	1602	Δ -	2607	421	 	
		1002	^	2057	421	[1	
MNVTGPQIRDGGVYRCTARNI VGSAEYQARI			l			1	VTWALDDEPIVEDGEHETMOVTMEDCTTIEU
253]		•		MNVTGPOIRDGGVYRCTARNI.VGSAEYOARI
FDDHITDSEALSRSSHVPTSHPRMLKRQPAIEL PLGGEYSSDVPRLSTGLSSSLLGYFSTLMTG AAFTNNIASSTILL PLGGEYSSDVPRLSTGLSSSLLGYFSTLMTG AAFTNNIASSTILL PLGGEYSSDVPRLSTGLSSSLLGYFSTLMTG AAFTNNIASSTILL PLGGEYSSDVPRLSTGLSSSLLGYFSTLMTG AAFTNNIASSTILL PLGGEYSSDVPRRLSTGLSSSLLGYFSTLMTG PRAGFEWAAPPGVGY PRAGF			l	1		ļ	
FDDHITDSEALSRSSHVPTSHPRMLKRQPARIE PLGGEYSSDVPRPLSTQLSSSLLGYPSTLMTG AAFTNNIASSTIIL GQIHSQDDPFIDQLGFGVAFQFQTFVACQEQ RYRGPWEAGPGVGY CQREQ RYRGPWEAGPGVGY CQREQ CQREGE CQREGE CQREGES CQREGE CQREGE CQREGES CQREGE	253	1603	A	2698	65	401	
PLGGEYSDVPRPLSTQLSSSLLGYFSTLMTG		1		} i		}	FDDHITDSEALSRSSHVFTSHPRMLKROPAIEL
254							PLGGEYSSDVPRPLSTQLSSSLLGYFSTLMTG
255							
255	254	1604	A	2699	438	301	
AQHLEVILKQKEQHYEEKIKVLDNQIKKDLA DKETLENMMQRHEEEAHEKGKLSEQKAMIN AMDSKIRSLEQRIVELSANKLAANSSLFTQR NMKAQEMISELRQOKFYLETQAGKLEAQN RKLEEQLEKISHQDHSDKNRLLELETRIREVS LEHEEQKLELKRQLTELQISLQERESQLTALQ AARAALSQLRQAKTELETTAEAEEGQALT VGLGSNIFRLLKASARMSVELALSILAHP VGLGSNIFRLLKASARMSVELALSILAHP TGLGSNIFRLLKASARMSVELALSILAHP TGLGSNIFRLLKASARMSVELALSILAHP TGLGSNIFRLKASARMSVELALSILAHP TGLGSNIFRLKASARMSVELALSILAHP TGLGSNIFRLKASARMSVELALSILAHP TGLGSNIFRLKASARMSVELALSILAHP TGLGSNIFRLKASARMSVELALSILAHP TGLGSNIFLLKASARMSVELALSILAHP TGLGGADPVAVMWDPRAARMDLTAYAE LLKEGNQVLKNONFSLAIRKYDEAIGLLQL TQWGVPPRDLAVILLONSNAFFSLGKWNEA FVAAKECLQWPPTYVKGYYRAGYSLLRIHQ PYEAARMFFEGLR TNPSTSTEDPRYSVGGDENIGTVTTLANLREF NPSLKGFSVGTGKETSPAFLNQAVAGGRAE DLPVQARRLVDLMKNDTRIHFQEDWKIITLFI GGNDL 258 1608 A 2709 1 1097 SVGARQGEARDRIRRFFFKGDLEVLQAQVERI MTRKELLTVYSSEDGSEFETIVLKALVKACG SSEASAYLDERLAVAMRVDLAQSELFRGDI QWRSFHLEASLMDALLANDRPEFVRLLISHGLS LGHFLTPMRLAQLYSAAPSNISLINLDQASH SAGTKAPALKGGAAELPPDVGHVLRMLLG KMCAPRYPSGGAWPHPQGGFGESMYLLISH ASTGTKAPALKGGAAELPPDVGHVLRMLLG KMCAPRYPSGGAWPHPQGGFGESMYLLISH ASTGTKAPALKGGAAELPPDVGHVLRMLLG KMCAPRYPSGGAWPHPQGFGESMYLLISH ASTGTKAPALKGGAAELPPDVGHVLRMLLG KMCAPRYPSGGAWPHPQGFGESMYLLISH ASTGTKAPALKGGAAELPPDVGHVLRMLLG KMCAPRYPSGGAWPHPQGFGESMYLLISH ASTGTKAPALKGGAAELPPDVGHVLRMLLG KMCAPRYPSGGAWPHPGGFGESMYLLISH ASTGTKAPALKGGAAELPPDVGHVLRMLLG KMCAPRYPSGGAWPHPGGFGESMYLLISH ASTGTKAPALKGGAGERPPDVGHVLRMLLG KMCAPRYPSGGAWPHPGGFGESMYLLISH ASTGTKAPALKGGAAELPPDVGHVLRMLLG KMCAPRYPSGGAWPHPGGFGESMYLLISH ASTGTKAPALKGGAAELPPDVGHVLRMLLG KMCAPRYPSGGAWPGGFSSPPEEEEEEEEGFGERA SCRRGLFRVRLTRVGLAAPSKASRGGEGDAA PKSPVREKSPKFRPRVSLSPKASRGGEGDAA PKSPVREKSPKFRPRVSLSPKASRGGEGDAA PKSPVREKSPKFRPRVSLSPKASRGGEGDAA PKSPVREKSPKFRPRVSLSPKASRGGEGDAA PKSPVREKSPKFRFRPVSLSPKASRGGEGDAA PKSPVREKSPKFRFRPVSLSPKASRGGEGDAA PKSPVREKSPKFRFRVSLSPKASRGGEGDAA PKSPVREKSPKFRFRVSLSPKASRGGEGDAA PKSPVREKSPKFRFRVSLSPKASRGGEGDAA PKSPVREKSPKFRFRVSLSPKASRGGEGDAA PKSPVREKSPKFRFRVSLSPKASRGGEGDAA PKSPVREKSPKFRFRVSLSPKASRGGEGDAA PKSPVGENGALLTRV DROPHTATATATATATATATATATATATATATATATATATATA	255	1605		2500			
DKETLENMMORHEEAAHEKGILSEQKAMIN AMDSKIRSLEGRIVEL SEANKLAANSSLFTOR NMKAQEEMISELRQQKFYLETQAGKLEAQN RKLEEQLEKISHQDHSDKNRILELETIKLEVS LEHEEQKELELKRQLTELQISLQERESQLTALQ AARAALESQLRQAKTELEFTTAEABEEIQALT VOLGSNIFRILKASARMSVELALSILAHP 256 1606 A 2701 2 405 FVGGPGADPPVAVMWDPRAARMDLTAYAE LLKESGNQVLKNGNFSLARKYDEAQILLQL YQWGVPPROLAVICKNSNAFFSLGKWNEA FVAAKECLQWDPTYVKGYYRAGYSLRLHQ PYEAARMFFEGLR PYEASASRPPGCFSGDGRFWLVSEGSRRGWD FNFSFSLDPRYSVGGDENIGTVTILANILREF NPSKGFSVGTGKEFSNAFLNQAVAGGRAE DLPVQARRLVDLMKNDTRIHFQEDWKIITLFI GGNDL 258 1608 A 2709 1 1097 SVGARQGEARDRIRRFFPKGDLEVLQAQVERI MTRKELLTVYSSEDGSEFETIVLKALVKACG SSEASAVLDELRLAVAWNRVDIAQSELFRGDI QWRSFHLEASLMDALLNDRPEFVRLLISHGLS LGHFLTPMRAAQLVSAAPSNSLIRNILDQASH SAGTKAPALKGGAAVLAVAWNRVDIAQSELFRGDI QWRSFHLEASLMDALLNDRPEFVRLLISHGLS LGHFLTPMRAAQLVSAAPSNSLIRNILDQASH SAGTKAPALKGGAAVEDERLAVAWNRVDIAQSH KATSPLSLDAGGLQGPGESMYLLSD KATSPLSLDAGGLQGPGESMYLLSD KATSPLSLDAGGLQGPGESMYLLSD KATSPLSLDAGGLQGPGESMYLLSD KATSPLSLDAGGLQAWGDHFQCGFGESMYLLSD KATSPLSLDAGGLQAWGDLLWALLLNRA QMAMYFWEMGSNAVSSALGACLLLRVMAR LEPDAEEAARRKDLAFFEGMOVDLFGECYR SSEVRARARLLLRRCPDVGHVLYMLG KMCAPRYPSGGAWDPHFQQFGESMYLLSD KATSPLSLDAGGLQAWGDLLLWALL NRA QMAMYFWEMGSNAVSSALGACLLLRVMAR LEPDAEEAARRKDLAFFEGMOVDLFGECYR SSEVRARARLLLRRCPDVGHVLYMLGI KMCAPRYPSGGAWDPHFQQFGESMYLLSD KATSPLSLDAGGLQAFSPSPEEEEEEEGGFGERA SCRRGLFRVRLTRVGLAAPSKASRGOGDDAA PKSPVERKSPKFPFRVSLSPKARSGSGDQEE GGLRVRLP 260 1610 A 2728 1 477 LLGGDLRYHLQNVHFTFEGTVKLYICELALA LEYLQRYHIIHRDIKPDNILLDEHGHVHITDFN IATVVKGAFRASSMAGTKPYMAPEVFQVYM DRGPGSYSYDDWWSLGITAYELLGAWRPYEI HSVTPIDEILNMFKVERVHYSSTWCKGMVAL LRK 261 1611 A 2730 3 547 LTITOFILVLYRYYRSPLVOIYEIEOHKIETWR	255	1000	A	2/00	I	842	LQNREDSSEGIRKKLVEAEELEEKHREAQVS
AMDSKIRSLEGRIVEL SEANKLAANSSLETQR NMKAQEMISELRQQKFYLETQAGKLEAQN RKLEEQLEKISHQDHSDKNRLLELETRLREVS LEHEEQKLEILKRQLTELQLSLQERESQLTALQ AARAALESQLRQAKTELEFTTAEAEBEQALT VGLGSNIFRILKASARMSVELALSILAHP 256 1606 A 2701 2 405 FVGGFGADPPVAVMWDPRAARMDLTAYAE LLKESGNQVLKNGNFSLAIRKYDEAGILLQL YQWGVPPRDLAVLLCNKSNAFFSLGKWNEA FVAAKECLQWDPTYVKGGYRAGYSLIRHQ PYEAARMFFEGLR 257 1607 A 2702 2 399 FVESASSRPFGCFSGDGRFWLVSEGSRRGWD FNPSSFLDPRYSVGGDENIGTVTTLANILREF NPSLKGFSVGTGKETSPNAFLNQAVAGGRAE DLPVQARRLVDLMKNDTRIHTQEDWKITLFI GGNDL 258 1608 A 2709 1 1097 SVGARQGEARDRIRRFFPKGDEVLQAQVERI MTRKELLTVYSSEDGSEFFFTIVLKALVKACG SSEASAYLDELRLAVAWNRVDLAQSELFRGDI QWRSFHLEASLMDALLNDRPFFVRLLISHGLS LGHFLTFMRLAQLYSAPSNSLIRNLDQASH SAGTKAPALKGGAAFLRPDVGHVLRMLLG KMCAPRYPSGGAWDPHFQGGFGESMYLLSD KATSPLSLDAGLQQAPWSDLLLWALLINRA QMAMYFWEMGSNAVSSALGACILLRVMAR LEPDAEAARKKDLAFKFEGMGVDLFGECYR SSEVRAARLLLRRCPLWGDATCLQLAMQAD ARAFFAQDGVQSLPTQKWWGDMARR VPLGRGGFLFFSNEGAKEGEKANIPKLMIPR GGFSQREMVTGERSPSPEEEEEEEGFGGRA SCRRGLFRVRLTRVGLAPSKASRGQEGDAA PKSPVREKSPKRFPRYVSLSPKARSGSGDQEE GGLRVRLP 260 1610 A 2728 1 477 LLGGDLRYHLQNVHFTEGTVKLYICELALA LEYLQRYHIIHRDKJDNILLDEHGHVHITDFN IATVVKGAFERASSMAGTKEPYMAPEVFQVYM DRGGFSYPVDWWSLGITAYBLLEGWRPYEI HSVTPIDEILNMFKVERVHYSSTWCKGMVAL LRK 261 1611 A 2730 3 547 LTITOFILVLYRYYRSPLVQIYEIEOHKIETWR							AQHLEVHLKQKEQHYEEKIKVLDNQIKKDLA
NMKAQEEMISELRQQKFYLETQAGKLEAON RKLEEQLEKISHQDHSDKNRLLELETRLREVS LEHEQKLEKISHQDHSDKNRLLELETRLREVS LEHEQKLELKQLTELQLSLQERESQLTALQ AARAALESQLRQAKTELETTTAEABEEIQALT VOLGSNIFRILKASARMSVELALSILAHP VOLGSNIFRILKASARMSVELALSILAHP VOLGSNIFRILKASARMSVELALSILAHP VOWGVPPDLAVLKNSNAFFSLGKWNEA FVGGPGADPFVAVMWDPRAARMDLTAYAE LLKESGNQVLKNGNFSLAIRKYDEAIQILLQL YQWGVPPDLAVLKNSNAFFSLGKWNEA FVAAKECLQWDPTYVKGYYRAGYSLLRLHQ PYEAARMFFGLR		ļ	ļ				
RKLEEGLEKISHQDHSDKNRLLELETRLREVS LEHEEQKLELKRQLTELQUSLQERESQLTALQ AARAALESQLRQAKTELEETTJAEAEEEIQALT VGLGSNIFRLIKASARMSVELALSILAHP VGLGSNIFRLIKASARMSVELALSILAHP FVGGGADPPVAMWDPRAARMDLTAYAE LLKESGNQVLKNGNFSLARKYDEAIQILLQL YQWGVPPRDLAVLLCNKSNAFFSLGKWNEA FVAAARECLQWPTYVKGYYRAG YSLLRLHQ PYEAARMFEGLR 257 1607 A 2702 2 399 FVESASSRPPGCFSGDGRFWLVSEGSRRGWD FNPSFSFLDPRYSVGGDENIGTVITLANILREF NPSLKGFSVGTGKETSPNAFLNQAVAGGRAE DLPVQARRLVDLMKNDTRIHFQEDWKITLFI GGNDL 258 1608 A 2709 1 1097 SVGARGGEARDRIRRFFFKGDLEVLQAQVERI MTRKELLTYVSSEDGSEFFETIVLKALVKACG SSEASAYLDELRLAVAWNRVDIAQSELFRGDI QWRSFHLEASLMDALLANDREFFVRLISHGLS UGWRSFHLEASLMDALLANDREFFVRLISHGLS KATSPLSLDAGLGQAFWSDLLLWALLLINRA QMAMYFVEMGSNAVSSALGACLLLRVMAR LEPDAEAARRADLAFKFEGMGVDLFGECYR SSEVRAARLLLRRCPLWGDATCLQLAMQAD ARAFFAQDGVQSLPTQKWGDMARR 259 1609 A 2721 1 403 VYLGAGPGLFSNEGKKEGEKANIPKLMLPR GGFSQREMVTGERSPSPEEEEEEEEGEGGRA SCRRGLFRVRLTRVGLAAPSKASRGGEGDAA PKSSVREKSPKFRPRVSLSPKARSGSGDQEE GGLRVRLP LLGGDLRYHLQQNVHFTEGTVKLYICELALA LEYLQRYHIHRDIKPDNILLDEHGHVHITDFN LATVVKGAERASSMAGTKYMAPEVFQUYM DRGFGYSYPVDWSLGITAYELLRGWRPYEI HSVTPIDEILNMFKVERVHYSSTWCKGMVAL LRK 261 1611 A 2730 3 547 LTITDFILVLYRYYRSPLVQIYEIEQHKIETWR							
LEHEEĞKLELKRÖLTELQISLQERESQLTALQ AARAALESQLRQAKTELEETTAEAEEEIQALT VOLGSNIFILLKASARMSVELALSILAHP			ļ				
AARAALESQLRQAKTELEETTAEAEEIQALT VGLGSNIFRLIKASARMSVELALSILAHP 256 1606 A 2701 2 405 FVGGFGADPPVAVMVDPRAARMDLTAYAE LLKESGNQVLKNGNFSLAIRKYDEAIQILLQL YQWGVPPRDLAVLLCNKSNAFFSLGKWNEA FVAAKECLQWDPTYVKGYYRAGYSLLRHQ PYEAARMFEGLR 257 1607 A 2702 2 399 FVESASSRPPGCFSGDGRFWLVSEGSRRGWD FNPSFSFLDPRYSVGDDENIGTVITLANILREF NPSLKGFSVGTGKETSPNAFLNQAVAGGRAE DLPVQARRLVDLMKNDTRIHFQEDWKIITLFI GGNDL 258 1608 A 2709 1 1097 SVGARQGEARDRIRRFPFKGDLEVLQAQVERI MTRKELLTVYSSEDGSEFFTIVLKALVKACG SSEASAYLDELRLAV AWNRVDIAQSELFRGDI QWRSFHLEASLMDALLNDRPEFVRLLISHGLS LGHFLTPMRLAQLYSAAPSNSLIRNLLDQASH SAGTKAPALKGGAAELRPPDVGHVLRMLLG KMCAPRYPSGGAWDPHPQGFGESMYLLSD KATSPLSLDAGLGQAPWSDLLLWALLLNRA QMAMYFWEMGSNAVSSALGACLLLRVMAR LEPDAEEAARRCDLAFKFEGMGVOLFGECYR SSEVRAARLLLRRCPLWGDATCLQLAMQAD ARAFFAQDGVQSLPTQKWWGDMARR 259 1609 A 2721 1 403 VYLGAGPGLFFSNEGAKEGEKANIPKLMLPR GGFSQREMVTGERSPSPEEEEEEEGFGERA SCRRGLFRVRLTRVGLAAPSKASRGQEGDAA PKSPVREKSPKFRFPRVSLSPKARSGSGDQEE GGLRVLP 260 1610 A 2728 1 477 LLGGDLRYHLQNVHFTEGTVKLYICELALA LEYLQRYHIHRDIKPDNILLDEHGHYHITDFN IATVVKGAERASSMAGTLLDEHGHYHITDFN IATVVKGAERASSMAGTLLDEHGHYHITDFN IATVVKGAERASSMAGTLLDEHGHYHITDFN IATVVKGAERASSMAGTLLDEHGHYHITDFN IATVVKGAERASSMAGTLLRGWRPYEI HSVTPIDEILNMFKVERVHYSSTWCKGMVAL LRK 261 1611 A 2730 3 547 LTITDFILVLYRYYRSPLVQIYEIEQHKIETWR	}						
VGLGSNIFRILKASARMSVELALSILAHF	1						AARAALESOLROAKTELEETTAEAEEEIOALT
256							VGLGSNIFRLLKASARMSVELALSILAHP
257 1607 A 2702 2 399 FVESASSRPGCFSGIGGRWLVSEGSRRGWD PYEAARMFFEGIR 258 1608 A 2709 1 1097 SVGARGEARDRIRRFFPKGDLEVLQAQVERI MTRKELLTVYSSEDGSEFETIVLKALVKACG SSEASAYLDELRLAVAWNRVDIAQSELFRGDI QWRSFHLEASLMDALLNDRPEFVRLLISHGIS LGHFLTRMLAQLYSAAPSNSLIRNLLDQASH SAGTKAPALKGGAAELRPDVGHVLRWLLG KMCAPRYPSGGAWDPHPGGFGESMYLLSD KATSPLSJDAGLGQAPWSDLLLWALLLNRA QMAMYFWEMGSNAVSSALGACLLLRVMAR LEPDAEEAARKDLAFKFEGMGVDLFGECYR SSEVRAARLLRRCPLWGDATCLQLAMQAD ARAFFAQDGVQSLPTQKWWGDMARR 259 1609 A 2721 1 403 VYLGAGPGLFFSNEGAKEGEKANIPKLMLPR GGFSQREWVTGERSPSPEEEEEEEEGFGERA SCRRGLFRVRLTRVGLAAPSKASRGQEGDAA PKSPVREKSPKFRFPRVSLSPKARSGSGDQEE GGLRVRLP 260 1610 A 2728 1 477 LLGGDLRYHLQNVHFTEGTVKLYICELALA LEYLQRYHIHRDIKPDNILLDEHGHVHTIDEN LATVVKGAERASSMAGTKPYMAPEVFQVYM DRGPGYSYPVDWWSLGITAYELLRGWRPYEI HSVTPIDEILNMFKVERVHYSSTWCKGMVAL LRK LTITDFILVLYRYYRSPLVQIYEIEQHKIETWR	256	1606	Α	2701	2	405	FVGGPGADPPVAVMWDPRAARMDLTAYAE
FVAAKECLQWDPTYVKGYYRAGYSLLRLHQ PYEARMFFEGLR	1			l			
257 1607 A 2702 2 399 FYESASSRPGGFSGDGRFWLVSEGSRRGWD FVESASSRPPGGFSGDGRFWLVSEGSRRGWD FNPSFSFLDPRYSVGGDENIGTVTTLANILREF NPSLKGFSVGTGKETSPNAFLNQAVAGGRAE DLPVQARRLVDLMKNDTRIHFQEDWKIITLFI GGNDL 258 1608 A 2709 1 1097 SVGARQGEARDRIRRFFFKGDLEVLQAQVERI MTRKELLTVYSSEDGSEEFETIVLKALVKACG SSEASAYLDELRLAVAWNRVDIAQSELFRGDI QWRSFHLEASLMDALLNDRPEFVRLISHGLS LGHFLTPMRLAQLYSAAPSNSLIRNLLDQASH SAGTKAPALKGGAAELRPPDVGHVLRMLLG KMCAPRYPSGGAWDPHPQGGFGESMYLLSD KATSPLSLDAGLQAPWSDLLLWALLLNRA QMAMYFWEMGSNAVSSALGACLLLRVMAR LEPDAEEAARRLDLAFKFEGMGVDLFGECYR SSEVRAARLLLRCPLWGDATCLQLAMQAD ARAFFAQGVQSLPTQKWWGDMARR 259 1609 A 2721 1 403 VYLGAGPGLFFSNEGAKEGEKANIPKLMLPR GGFSQREMVTGERSPSPEEEEEEEEGFGGRA SCRRGLFRVRLTNVGLAAPSKASRGQEGDAA PKSPVREKSPKFRFPRVSLSPKARSGSGDQEE GGLRVRLP 260 1610 A 2728 1 477 LLGGDLRYHLQONVHFTEGTVKLYICELALA LEYLQRYHIIHRDIKPDNILLDEHGHVHITDFN IATVVKGAERASSMAGTKPYMAPEVFQVYM DRGPGYSYPVDWWSLGITAYELLRGWRPYEI HSVTPIDEILNMFKVERVHYSSTWCKGMVAL LRK 261 1611 A 2730 3 547 LTITDFILVLYRYYRSPLVQIYEIEOHKIETWR	1	 					
257		;					
FNPSFSFLDPRYSVGGDENIGTVTTLANILREF NPSLKGFSVGTGKETSPNAFLNQAVAGGRAE DLPVQARRLVDLMKNDTRIHFQEDWKIITLFI GGNDL	257	1607		2702	2	200	
258 1608 A 2709 1 1097 SVGARQGEARDRIRRFFPKGDLEVLQAQVERI GGNDL WARSHLLEASLMDALLANDRPEFVRLLISHGLS LGHFLTPMRLAQLYSAAPSNSLIRNLLDQASH SAGTKAPALKGGAAELRPPDVGHVLRMLLG KMCAPRYPSGGAWDPHPQQGFGESMYLLSD KATSPLSLDAGLGQAPWSDLLLWALLINRA QMAMYFWEMGSNAVSSALGACLLLRVMAR LEPDAEEAARRKDLAFKFEGMGVDLFGECYR SSEVRAARLLRRCPLWGDATCLQLAMQAD ARAFFAQDGVQSLPTQKWWGDMARR 259 1609 A 2721 1 403 VYLGAGPGLFFSNEGAKEGEKANIPKLMLPR GGFSQREMVTGERSPSPEEEEEEGFGERA SCRGLFRVRLTRVGLAAPSKASRGQEGDAA PKSPVREKSPKFFPRVSLSPKARSGSDQEE GGLRVRLP 260 1610 A 2728 1 477 LLGGDLRYHLQQNVHFTEGTVKLYICELALA LEYLQRYHIHRDIKPDNILLDEHGHVHITDFN LATVVKGAERASSMAGTKPYMAPEVFQVYM DRGPGYSYPVDWWSLGITAYELLRGWRPYEI HSVTPIDEILNMFKVERVHYSSTWCKGMVAL LRK 261 1611 A 2730 3 547 LTITDFILVLYRYYRSPLVQIYEIEQHKIETWR	237	1007	Λ.	2/02		399	
DLPVQARRLVDLMKNDTRIHFQEDWKIITLFI GGNDL 258 1608 A 2709 1 1097 SVGARQGEARDRIRRFFFKGDLEVLQAQVERI MTRKELLTYYSSEDGSEEFETIVLKALVKACG SSEASAYLDELRLAVAWNRVDIAQSELFRGDI QWRSFHLEASLMDALLNDRPEFVRLLISHGLS LGHFLTPMRLAQLYSAAPSNSLIRNLLDQASH SAGTKAPALKGGAAELRPPDVGHVLRMLLG KMCAPRYPSGGAWDPHPQQGFGESMYLLSD KATSPLSLDAGLGQAPWSDLLLWALLLNRA QMAMYFWEMGSNAVSSALGACLLLRVMAR LEPPDAEEAARRKDLAFKFEGMGVDLFGECYR SSEVRAARLLRRCPLWGDATCLQLAMQAD ARAFFAQDGVQSLPTQKWWGDMARR 259 1609 A 2721 1 403 VYLGAGPGLFSNEGAKEGEKANIPKLMLPR GGFSQREMVTGERSPSPEEEEEEEGFGERA SCRRGLFRVRLTRVGLAAPSKASRGQEGDAA PKSPVREKSPKFFPRVSLSPKARSGSGDQEE GGLRVRLP 260 1610 A 2728 1 477 LLGGDLRYHLQNVHFTEGTVKLYICELALA LEYLORYHIHRDIKPDNILLDEHGHVHITDFN IATVVKGAERASSMAGTKPYMAPEVFQVYM DRGPGYSYPVDWWSLGITAYELLRGWRPYEI HSVTPIDELNMFKVERVHYSSTWCKGMVAL LRK 261 1611 A 2730 3 547 LTITDFILVLYRYYRSPLVQIYEIEQHKIETWR	 						
GGNDL GGNDL SVGARQGEARDRIRRFFPKGDLEVLQAQVERI MTRKELLTVYSSEDGSEFFETIVLKALVKACG SSEASAYLDELRLAVAWNRVDIAQSELFRGDI QWRSFHLEASLMDALLNDRPEFVRLLISHGLS LGHFLTPMRLAQLYSAAPSNSLIRNLLDQASH SAGTKAPALKGGAAELRPPDVGHVLRMLLG KMCAPRYPSGGAWDPHPCQGFGESMYLLSD KATSPLSLDAGLGQAPWSDLLLWALLLNRA QMAMYFWEMGSNAVSSALGACLLLRVMAR LEPDAEEAARRKDLAFKFEGMGVDLFGECYR SSEVRAARLLLRRCPLWGDATCLQLAMQAD ARAFFAQDGVQSLPTQKWWGDMARR CGFSQREMVTGERSPSPEEEEEEEEFGGERA SCRRGLFRVRLTRVGLAAPSKASRGQEGDAA PKSPVRLKSPKFFFPRVSLSPKARSGSGDQEE GGLRVRLP GGFSQREMVTGERSPSPEEEEEEEEFGGERA SCRRGLFRVRLTRVGLAAPSKASRGQEGDAA PKSPVRLKSPKFFPRVSLSPKARSGSGDQEE GGLRVRLP LEGGDLRYHLQQNVHFTEGTVKLYICELALA LEYLQRYHIHRDIKPDNILLDEHGHVHITDFN LATVVKGAERASSMAGTKPYMAPEVFQVYM DRGPGYSYPVDWSLGITAYELLRGWRPYEI HSVTPIDEILNMFKVERVHYSSTWCKGMVAL LRK LTITDFILVLYRYYRSPLVQIYEIEQHKIETWR LTITDFILVLYRYYRSPLVQIYEIEQHKIETWR							DLPVOARRI.VDLMKNDTRIHFOEDWKIITI FI
258 1608 A 2709 1 1097 SVGARQGEARDRIRRFFPKGDLEVLQAQVERI MTRKELLTVYSSEDGSEFFTIVLKALVKACG SSEASAYLDELRLAVAWNRVDIAQSELFRGDI QWRSFHLEASLMDALLNDRPEFVRLLISHGLS LGHFLTPMRLAQLYSAAPSNSLIRNLLDQASH SAGTKAPALKGGAAELRPPDVGHVLRMLLG KMCAPRYPSGGAWDPHPGQGFGESMYLLSD KATSPLSLDAGLGQAPWSDLLLWALLLNRA QMAMYFWEMGSNAVSSALGACLLLRVMAR LEPDAEEAARRKDLAFKFEGMGVDLFGECYR SSEVRAARLLLRRCPLWGDATCLQLAMQAD ARAFFAQDGVQSLPTQKWWGDMARR 259 1609 A 2721 1 403 VYLGAGPGLFFSNEGAKEGEKANIPKLMLPR GGFSQREMVTGERSPSPEEEEEEEEGFGERA SCRRGLFRVRLTRVGLAAPSKASRGQEGDAA PKSPVREKSPKFRFPRVSLSPKARSGSGDQEE GGLRVLP 260 1610 A 2728 1 477 LLGGDLRYHLQQNVHFTEGTVKLYICELALA LEYLQRYHIIHRDIKPDNILLDEHGHVHTTDFN IATVVKGAERASSMAGTKPYMAPEVFQVYM DRGPGYSYPVDWWSLGITAYELLRGWRPYEI HSVTPIDEILNMFKVERVHYSSTWCKGMVAL LRK 261 1611 A 2730 3 547 LTITDFILVLYRYYRSPLVQIYEIEQHKIETWR							
MTRKELLTVYSSEDGSEEFETIVLKALVKACG SSEASAYLDELRLAVAWNRVDIAQSELFRGDI QWRSFHLEASLMDALLNDRPEFVRLLISHGLS LGHFLTPMRLAQLYSAAPSNSLIRNLLDQASH SAGTKAPALKGGAAELRPPDVGHVLRMLLG KMCAPRYPSGGAWDPHPQQGFGESMYLLSD KATSPLSLDAGLGQAPWSDLLLWALLLNRA QMAMYFWEMGSNAVSSALGACLLLRVMAR LEPDAEEAARRKDLAFKFEGMGVDLFGECYR SSEVRAARLLLRCPL WGDATCLQLAMQAD ARAFFAQDGVQSLPTQKWWGDMARR 259 1609 A 2721 1 403 VYLGAGPGLFFSNEGAKEGEKANIPKLMLPR GGFSQREMVTGERSPSPEEEEEEEGFGERA SCRRGLFRVRLTRVGLAAPSKASRGQEGDAA PKSPVREKSPKFFPRVSLSPKARSGSGDQEE GGLRVRLP 260 1610 A 2728 1 477 LLGGDLRYHLQQNVHFTEGTVKLYICELALA LEYLQRYHIIHRDIKPDNILLDEHGHVHITDFN IATVVKGAERASSMAGTKPYMAPEVFQVYM DRGPYSYPVDWWSLGITAYELLRGWRPYEI HSVTPIDEILNMFKVERVHYSSTWCKGMVAL LRK 261 1611 A 2730 3 547 LTITDFILVLYRYYRSPLVQIYEIEQHKIETWR	258	1608	A	2709	1	1097	
SSEASAYLDELRLAVAWNRVDIAQSELFRGDI QWRSFHLEASLMDALLNDRPEFVRLLISHGLS LGHFLTPMRLAQLYSAAPSNSLIRNILDQASH SAGTKAPALKGGAAELRPPDVGHVLRMLLG KMCAPRYPSGGAWDPHPQGFGESMYLLSD KATSPLSLDAGLGQAPWSDLLLWALLLNRA QMAMYFWEMGSNAVSSALGACLLLRVMAR LEPDAEEAARRKDLAFKFEGMGVDLFGECYR SSEVRAARLLLRCPLWGDATCLQLAMQAD ARAFFAQDGVQSLPTQKWWGDMARR 259 1609 A 2721 1 403 VYLGAGPGLFFSNEGAKEGEKANIPKLMLPR GGFSQREMVTGERSPSPEEEEEEEEGFGERA SCRRGLFRVRLTRVGLAAPSKASRGQEGDAA PKSPVREKSPKFRFPRVSLSPKARSGSGDQEE GGLRVRLP 260 1610 A 2728 1 477 LLGGDLRYHLQQNVHFTEGTVKLYICELALA LEYLQRYHIHRDIKPDNILLDEHGHVHITDFN IATVVKGAERASSMAGTKPYMAPEVFQVYM DRGPGYSYPVDWWSLGITAYELLRGWRPYEI HSVTPIDEILNMFKVERVHYSSTWCKGMVAL LRK 261 1611 A 2730 3 547 LTITDFILVLYRYYRSPLVQIYEIEQHKIETWR	}						MTRKELLTVYSSEDGSEEFETIVLKALVKACG
LGHFLTPMRLAQLYSAAPSNSLIRNLLDQASH SAGTKAPALKGGAAELRPPDVGHVLRMLLG KMCAPRYPSGGAWDPHPGQGFGESMYLLSD KATSPLSLDAGLGQAPWSDLLLWALLLNRA QMAMYFWEMGSNAVSSALGACLLLRVMAR LEPDAEEAARRKDLAFKFEGMGVDLFGECYR SSEVRAARLLLRRCPLWGDATCLQLAMQAD ARAFFAQDGVQSLPTQKWWGDMARR 259 1609 A 2721 1 403 VYLGAGPGLFFSNEGAKEGEKANIPKLMLPR GGFSQREMVTGERSPSPEEEEEEEEGFGERA SCRRGLFRVRLTRVGLAAPSKASRGQEGDAA PKSPVREKSPKFRFPRVSLSPKARSGSGDQEE GGLRVRLP 260 1610 A 2728 1 477 LLGGDLRYHLQQNVHFTEGTVKLYICELALA LEYLQRYHIHRDIKPDNILLDEHGHVHITDFN IATVVKGAERASSMAGTKPYMAPEVFQVYM DRGPGYSYPVDWWSLGITAYELLRGWRPYEI HSVTPIDEILNMFKVERVHYSSTWCKGMVAL LRK 261 1611 A 2730 3 547 LTITDFILVLYRYYRSPLVQIYEIEQHKIETWR							SSEASAYLDELRLAVAWNRVDIAQSELFRGDI
SAGTKAPALKGGAAELRPPDVGHVLRMLLG KMCAPRYPSGGAWDPHPGQGFGESMYLLSD KATSPLSLDAGLGQAPWSDLLLWALLLNRA QMAMYFWEMGSNAVSSALGACLILRVMAR LEPDAEEAARRKDLAFKFEGMGVDLFGECYR SSEVRAARLLLRCPLWGDATCLQLAMQAD ARAFFAQDGVQSLPTQKWWGDMARR 259 1609 A 2721 1 403 VYLGAGPGLFFSNEGAKEGEKANIPKLMLPR GGFSQREMVTGERSPSPEEEEEEEEGFGERA SCRRGLFRVRLTRVGLAAPSKASRGQEGDAA PKSPVREKSPKFRFPRVSLSPKARSGSGDQEE GGLRVRLP 260 1610 A 2728 1 477 LLGGDLRYHLQQNVHFTEGTVKLYICELALA LEYLQRYHIHRDIKPDNILLDEHGHVHITDFN IATVVKGAERASSMAGTKPYMAPEVFQVYM DRGPGYSYPVDWWSLGITAYELLRGWRPYEI HSVTPIDEILNMFKVERVHYSSTWCKGMVAL LRK 261 1611 A 2730 3 547 LTITDFILVLYRYYRSPLVQIYEIEQHKIETWR							
KMCAPRYPSGGAWDPHPGQGFGESMYLLSD KATSPLSLDAGLGQAPWSDLLLWALLLNRA QMAMYFWEMGSNAVSSALGACLLLRVMAR LEPDAEEAARRKDLAFKFEGMGVDLFGECYR SSEVRAARLLLRRCPLWGDATCLQLAMQAD ARAFFAQDGVQSLPTQKWWGDMARR VYLGAGPGLFFSNEGAKEGEKANIPKLMLPR GGFSQREMVTGERSPSPEEEEEEEEEGFGERA SCRRGLFRVRLTRVGLAAPSKASRGQEGDAA PKSPVREKSPKFRFPRVSLSPKARSGSGDQEE GGLRVRLP LLGGDLRYHLQQNVHFTEGTVKLYICELALA LEYLQRYHIHRDIKPDNILLDEHGHVHITDFN IATVVKGAERASSMAGTKPYMAPEVFQVYM DRGPGYSYPVDWWSLGITAYELLRGWRPYEI HSVTPIDEILNMFKVERVHYSSTWCKGMVAL LRK 261 1611 A 2730 3 547 LTITDFILVLYRYYRSPLVQIYEIEQHKIETWR							
KATSPLSLDAGLGQAPWSDLLLWALLNRA QMAMYFWEMGSNAVSSALGACLLLRVMAR LEPDAEEAARRKDLAFKFEGMGVDLFGECYR SSEVRAARLLLRRCPLWGDATCLQLAMQAD ARAFFAQDGVQSLPTQKWWGDMARR VYLGAGPGLFFSNEGAKEGEKANIPKLMLPR GGFSQREMVTGERSPSPEEEEEEEEGFGERA SCRRGLFRVRLTRVGLAAPSKASRGQEGDAA PKSPVREKSPKFRFPRVSLSPKARSGSDQEE GGLRVRLP LLGGDLRYHLQQNVHFTEGTVKLYICELALA LEYLQRYHIHRDIKPDNILLDEHGHVHITDFN IATVVKGAERASSMAGTKPYMAPEVFQVYM DRGPGYSYPVDWWSLGITAYELLRGWRPYEI HSVTPIDEILNMFKVERVHYSSTWCKGMVAL LRK 261 1611 A 2730 3 547 LTITDFILVLYRYYRSPLVQIYEIEQHKIETWR							
QMAMYFWEMGSNAVSSALGACLLLRVMAR LEPDAEEAARRKDLAFKFEGMGVDLFGECYR SSEVRAARLLLRRCPLWGDATCLQLAMQAD ARAFFAQDGVQSLPTQKWWGDMARR 259 1609 A 2721 1 403 VYLGAGPGLFFSNEGAKEGEKANIPKLMLPR GGFSQREMVTGERSPSPEEEEEEEEGFGERA SCRRGLFRVRLTRVGLAAPSKASRGQEGDAA PKSPVREKSPKFRFPRVSLSPKARSGSGDQEE GGLRVRLP 260 1610 A 2728 1 477 LLGGDLRYHLQQNVHFTEGTVKLYICELALA LEYLQRYHIIHRDIKPDNILLDEHGHVHITDFN LATVVKGAERASSMAGTKPYMAPEVFQVYM DRGPGYSYPVDWWSLGITAYELLRGWRPYEI HSVTPIDEILNMFKVERVHYSSTWCKGMVAL LRK 261 1611 A 2730 3 547 LTITDFILVLYRYYRSPLVQIYEIEQHKIETWR			Į		l		
LEPDAEEAARRKDLAFKFEGMGVDLFGECYR SSEVRAARLLLRRCPLWGDATCLQLAMQAD ARAFFAQDGVQSLPTQKWWGDMARR 259 1609 A 2721 1 403 VYLGAGPGLFFSNEGAKEGEKANIPKLMLPR GGFSQREMVTGERSPSPEEEEEEEEGFGERA SCRRGLFRVRLTRVGLAAPSKASRGQEGDAA PKSPVREKSPKFRFPRVSLSPKARSGSGDQEE GGLRVRLP 260 1610 A 2728 1 477 LLGGDLRYHLQQNVHFTEGTVKLYICELALA LEYLQRYHIIHRDIKPDNILLDEHGHVHITDFN LATVVKGAERASSMAGTKPYMAPEVFQVYM DRGPGYSYPVDWWSLGITAYELLRGWRPYEI HSVTPIDEILNMFKVERVHYSSTWCKGMVAL LRK 261 1611 A 2730 3 547 LTITDFILVLYRYYRSPLVQIYEIEQHKIETWR	 				J		
SSEVRAARLLRRCPLWGDATCLQLAMQAD ARAFFAQDGVQSLPTQKWWGDMARR 259 1609 A 2721 1 403 VYLGAGPGLFFSNEGAKEGEKANIPKLMLPR GGFSQREMVTGERSPSPEEEEEEEEGFGERA SCRRGLFRVRLTRVGLAAPSKASRGQEGDAA PKSPVREKSPKFRFPRVSLSPKARSGSGDQEE GGLRVRLP 260 1610 A 2728 1 477 LLGGDLRYHLQQNVHFTEGTVKLYICELALA LEYLQRYHIIHRDIKPDNILLDEHGHVHITDFN LATVVKGAERASSMAGTKPYMAPEVFQVYM DRGPGYSYPVDWWSLGITAYELLRGWRPYEI HSVTPIDEILNMFKVERVHYSSTWCKGMVAL LRK 261 1611 A 2730 3 547 LTITDFILVLYRYYRSPLVQIYEIEQHKIETWR			- 1	.	l		
259 1609 A 2721 1 403 VYLGAGPGLFFSNEGAKEGEKANIPKLMLPR GGFSQREMVTGERSPSPEEEEEEEGFGERA SCRRGLFRVRLTRVGLAAPSKASRGQEGDAA PKSPVREKSPKFRFPRVSLSPKARSGSGDQEE GGLRVRLP 260 1610 A 2728 1 477 LLGGDLRYHLQQNVHFTEGTVKLYICELALA LEYLQRYHIIHRDIKPDNILLDEHGHVHITDFN IATVVKGAERASSMAGTKPYMAPEVFQVYM DRGPGYSYPVDWWSLGITAYELLRGWRPYEI HSVTPIDEILNMFKVERVHYSSTWCKGMVAL LRK 261 1611 A 2730 3 547 LTITDFILVLYRYYRSPLVQIYEIEQHKIETWR							
259 1609 A 2721 1 403 VYLGAGPGLFFSNEGAKEGEKANIPKLMLPR GGFSQREMVTGERSPSPEEEEEEEGFGERA SCRRGLFRVRLTRVGLAAPSKASRGQEGDAA PKSPVREKSPKFRFPRVSLSPKARSGSGDQEE GGLRVRLP 260 1610 A 2728 1 477 LLGGDLRYHLQQNVHFTEGTVKLYICELALA LEYLQRYHIIHRDIKPDNILLDEHGHVHITDFN IATVVKGAERASSMAGTKPYMAPEVFQVYM DRGPGYSYPVDWWSLGITAYELLRGWRPYEI HSVTPIDEILNMFKVERVHYSSTWCKGMVAL LRK 261 1611 A 2730 3 547 LTITDFILVLYRYYRSPLVQIYEIEQHKIETWR	L						ARAFFAQDGVQSLPTQKWWGDMARR
GGFSQREMVTGERSPSPEEEEEEEEGFGERA SCRRGLFRVRLTRVGLAAPSKASRGQEGDAA PKSPVREKSPKFRFPRVSLSPKARSGSGDQEE GGLRVRLP LUGGDLRYHLQQNVHFTEGTVKLYICELALA LEYLQRYHIHRDIKPDNILLDEHGHVHITDFN LATVVKGAERASSMAGTKPYMAPEVFQVYM DRGPGYSYPVDWWSLGITAYELLRGWRPYEI HSVTPIDEILNMFKVERVHYSSTWCKGMVAL LRK LTITDFILVLYRYYRSPLVQIYEIEQHKIETWR	259	1609	A	2721	1	403	VYLGAGPGLFFSNEGAKEGEKANIPKLMLPR
PKSPVREKSPKFRFPRVSLSPKARSGSGDQEE GGLRVRLP 260 1610 A 2728 1 477 LLGGDLRYHLQQNVHFTEGTVKLYICELALA LEYLQRYHIHRDIKPDNILLDEHGHVHITDFN IATVVKGAERASSMAGTKPYMAPEVFQVYM DRGPGYSYPVDWWSLGITAYELLRGWRPYEI HSVTPIDEILNMFKVERVHYSSTWCKGMVAL LRK 261 1611 A 2730 3 547 LTITDFILVLYRYYRSPLVQIYEIEQHKIETWR	†		l	J	. 1	ł	GGFSQREMVTGERSPSPEEEEEEEEGFGERA
260 1610 A 2728 1 477 LLGGDLRYHLQQNVHFTEGTVKLYICELALA LEYLQRYHIIHRDIKPDNILLDEHGHVHITDFN IATVVKGAERASSMAGTKPYMAPEVFQVYM DRGPGYSYPVDWWSLGITAYELLRGWRPYEI HSVTPIDEILNMFKVERVHYSSTWCKGMVAL LRK 261 1611 A 2730 3 547 LTITDFILVLYRYYRSPLVQIYEIEQHKIETWR				į	ļ		SCRRGLFRVRLTRVGLAAPSKASRGQEGDAA
260 1610 A 2728 1 477 LLGGDLRYHLQQNVHFTEGTVKLYICELALA LEYLQRYHIIHRDIKPDNILLDEHGHVHITDFN IATVVKGAERASSMAGTKPYMAPEVFQVYM DRGPGYSYPVDWWSLGITAYELLRGWRPYEI HSVTPIDEILNMFKVERVHYSSTWCKGMVAL LRK 261 1611 A 2730 3 547 LTITDFILVLYRYYRSPLVQIYEIEQHKIETWR) l		ł	ŀ	- 1	ł	
LEYLQRYHIIHRDIKPDNILLDEHGHVHITDFN IATVVKGAERASSMAGTKPYMAPEVFQVYM DRGPGYSYPVDWWSLGITAYELLRGWRPYEI HSVTPIDEILNMFKVERVHYSSTWCKGMVAL LRK 261 1611 A 2730 3 547 LTITDFILVLYRYYRSPLVQIYEIEQHKIETWR	360	1610		2720	,	455	
IATVVKGAERASSMAGTKPYMAPEVFQVYM DRGPGYSYPVDWWSLGITAYELLRGWRPYEI HSVTPIDEILNMFKVERVHYSSTWCKGMVAL LRK 261 1611 A 2730 3 547 LTITDFILVLYRYYRSPLVQIYEIEQHKIETWR	200	1010	A	2/28	1	477	
DRGPGYSYPVDWWSLGITAYELLRGWRPYEI HSVTPIDEILNMFKVERVHYSSTWCKGMVAL LRK 261 1611 A 2730 3 547 LTITDFILVLYRYYRSPLVQIYEIEQHKIETWR				1			
HSVTPIDEILNMFKVERVHYSSTWCKGMVAL LRK 261 1611 A 2730 3 547 LTITDFILVLYRYYRSPLVQIYEIEQHKIETWR		1	ł	1	İ		`
261 1611 A 2730 3 547 LTITDFILVLYRYYRSPLVQIYEIEQHKIETWR		ĺ		I	1	i	
261 1611 A 2730 3 547 LTITDFILVLYRYYRSPLVQIYEIEQHKIETWR			1	}		ł	
EIYLOGCFKPLVSISPNDSLFEAVYTLIKNRIH	261	1611	A	2730	3	547	
							EIYLQGCFKPLVSISPNDSLFEAVYTLIKNRIH

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:			Amino acid sequence (A=Alanine C=Cysteine,
1	peptide	noa	1	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-		1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1	ļ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1	ł			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	ł.	i	ļ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	ł	l	ł	peptide	Joquomoo	/=possible nucleotide deletion, \=possible
1			i	sequence		nucleotide insertion
<u> </u>		<u> </u>		sequence		nucleotide insertion
1			[RLPVLDPVSGNVLHILTHKRLLKFLHIFGSLLP
1		į	ĺ			RPSFLYRTIQDLGIGTFRDLAVVLETAPILTAL
1		1	1		!	DIFVDRRVSALAVVNECGTHPQDERLGLGW
			ļ			GLGEPGSEERLFPAAITSR
262	1612	A	2733	3	431	GPEFPGSAKLVFLDLSYNNLTQLGAGAFRSA
1	i	ł			1	GRLVKLSLANNNLVGVHEDAFETLESLQVLE
}	ļ					I VIDABIL DEL CUL AL ALL DAL DOLDA DE DESTOVE
		ļ				LNDNNLRSLSVAALAALPALRSLRLDGNPWL
						CDCDFAHLFSWIQENASKLPKGLDEIQCSLPM
<u> </u>						ESRRISLRACRRPASRV
263	1613	A _	2736	2	343	PARISGVDPPVRKATKGGENCSFEDNKNWQF
1	}					LWGLNGNFNFFKEPWGGRNNHAKGFRTTW
]	i			ARSSSQNNRTFQNNRNFLRLQRDSQKKGQFA
i .						RLISPLVNLPQSPGGLEFQYQAT
264	1614	A	2738	2	245	DAM KOLDBOODDOO
204	1014	^	2/30	4	245	RAMLKCLREGQPPPSYNWTRLDGPLPSGVRV
, l						DGDTLGFPPLTTEHSGIYVRHDTNEFSSRDSH
					L	DTVDVLDPPEDSGKQVDL
265	1615	Α	2752	2	388	AAGDAPLRSLEQANRTRFPFFSDVKGDHRLV
1 1		1	١ .			LAAVETTVLVLIFAVSLLGNVCALVLVARRR
1		1				RRGATACLVLNLFCADLLFISAIPLVLAVRWT
		1				CAME CONTACT TO THE CONTEST OF THE C
1 1			i			EAWLLGPVACHLLFYVMTLSGSVTILTLAAV
<u> </u>						SLER
266	1616	Α]	2755	192	1	AFREVGGYWGLLCEHLYAIPSKTSEGNWTAK
						LQGYLPLQDAFHIFQDPLTGDLPWPELILGLP
!						l v
267	1617	Α	2760	434	714	ASRLEKQNSTPESDYDNTPNDMEPDGMGYM
1 1					, - ,	HRTSVPGEGLPRARDLAGLGOOKOFTTHTPF
			i			I VEOTUVCI V DOGIDORIVATOI CIO COMPINATOR
268	1618	Α	2762	1	105	LYFQTHKGLKDSSIRSEVTCLGISQCWRKGFF
200	1019	A	2/02	1	405	IACTFCGQDEWSPERSTRCFRRSRFLAWGEP
						AVLLLLLLSLALGLVLAALGLFVHHRDSPL
1 1			1			VQASGGPLACFGLVCLGLVCLSVLLFPGQPSP
		İ				ARCLAQQPLSHLPLTGCLSTLFLOAAEIFVESE
		.	- 1		,	LPLSWAE
269	1619	A	2772	3	243	TRPAEKIQYLVLFFVMSHPSQAYDKLSLSDHL
				-	2.5	LIAVLNLLRREVSEHGRHLQQYFNLFVMYAN
J	ļ	j	,			
270	1600		2700		106	LSKNLSFSEFCFDVSY
270	1620	Α	2789	1	486	ELQSQQACTHTKETEQLRSQLQTLKQQHQQA
	ſ	1	!		ŀ	VEQIAKAEETHSSLSQELQARLQTVTREKEEL
j !	1	j	Ì		l	LQLSIERGKVLQNKQAEICQLEEKLEIANEDR
j j	j				l	KHALERFEQEAVAVDSNLRVRELQRKVDGIQ
		- 1	[ĺ	KAYDELRLQSEAFKKHSLDLLSKERELNGKL
						RHLSP
271	1621	A	2795	1	568	
	.021	^	2173	•	200	KEKRVTVQLPTESIQKNQEDKLKMVPRKQRE
1		1		İ		FSGSDRGKLPGSEEKNQGPSMIGRKEERLITE
l	İ			-		RKHEHLKNKSAPKVVKQKVIDAHLDSQTQN
i	l	- 1	l	l		FQQTQIQTAESKAEHKKLPQPYNSLQEEKCLE
		i		1		VKGIQEKQVFSNTKDSKQEITQNKSFFSSVKE
ļ j	}	J		I	J	SQRDDGKGALNIVEFLRKREELHQILSTVKOP
272	1622	A	2797	8	523	KCMQGKYAGAMESEPCVCTEADFDCDYGYE
		- 1		~	1	
1	į	ŀ	1	1	1	RHSNGQCLPAFWFNPSSLSKDCSLGQSYLNST
- 1	į			ļ	i	GYRKVVSNNCTDGVREQYTAKPQKCPGKAP
1	İ	- 1				RGLRIVTADGKLTAEQGHNVTLMVQLEEGD
1	ĺ	ĺ	- {	1	1	VQRTLIQVDFGDGIAVSYVNLSSMEDGIXHV
- 1	-	1	- 1	İ	l	YONXGIXRXTVQVDNSLGS
273	1623	A	2801	72	395	HPSRSNVGPRQLTVWNTSNLSHDNRRKYIFS
		[· ~		
		- 1	1	ļ	1	DEEGQNQLGIRIHQDIPLPPRRRELPALRTTNG
•	- (- 1	1	}	ĺ	KADSLNVSRNSVMQELSELEKQIQVIRQELQL
	لحبيب					AVSRKTELEEYH
274	1624	Α	2805	168	320	ILWLYFETGTWVYPVFAKLSLLGLAALFSLRE
		- }	ļ		j	IFIARNGVVGETLTHCKRV
						

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ł	09/496 914	correspondi	to last amino acid residue	M=Methionine, N=Asparagine, P=Proline,
uence	Ì		914	ng to first amino acid	of peptide	Q=Glutamine, R=Arginine, S=Serine,
1		[residue of	sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
]				peptide	Sequence	/=possible nucleotide deletion, \=possible
j	J		ļ	sequence		nucleotide insertion
275	1625	A	2812	208	321	GSLATCQLSEPLLWFILRVLDTSDALKAFHD
		}				MGKIIFO
276	1626	A	2813	41	266	AGRSLHGAGDRAWVGISPTDWSPKVVELCK
İ		ļ				KYQQQTVVAIDLAGDETIPGSSLLPGHVQAY
						QVGPVRRNGEAGPG
277	1627	Α	2817	3	410	VLQERLDNFQRKCIQLASSTEGKVDKLLMRN
1.			i i		i	LFISYLHTPKHKQHEVLQAMGSILGITGEEME
Í		Ì	1		}	PLFQEEHGTATRWMTGWLEGGSKSVPKTPL
						GLNQQPALNGSFSELFVKFLKTESLSSTLPTX
278	1628	A	2821	238	467	LPPHNSPGKIK
2/8	1028	A	2821	238	457	GLSGPSCSCPHSPLPTIISRAQLETALKWRNYE
			!			VKLRLLLHLEELQMEHDIRHYDLESVPMTWD PVDQNPRLV
279	1629	Ā	2822	342	1	PLIPANLPAHSNPLQPLPSLPHPFLPATHKFPT
2.7	1029	Λ.	2022	342	1	TPPTFSSVPPPLPSLSSILHHSPLHSELNPHLOS
1			1			CRLPSRPSVSRELPPQSGPASSVPLAPTPLPDS
						VPSQRHPTXPPPAS
280	1630	A	2825	307	77	PSMVWSYHWGVKQKRLALCVFSFEEGGRRK
1						CGQYWPLEKDSRIRFGFLTVTNLTGAVGEPG
						VAFQCDGQRRREPTC
281	1631	A	2827	81	381	KMGTAVWVPKEKEKRDKASQEGGDVLGAR
1		•				QDCTPSLKSLVATGNLLDLEETAKAPLSTVSA
				,		NTTNMDEVPRPQALSGSSVVWVSGCVASRS
	1600					VILSLTSG
282	1632	A	2830	471	160	KLPXDKYELEPSPLTQYILERKSPHTCWQVFV
						TSSGKYNELGYPFGYLKASTTLTCVNLFVMP
		:				YNYPVLLPLLDDLFKVHKLKPNLKWRQAFDS
283	1633	A	2835	462	148	YLKTLPPYYL VSPALSLTPTIFSYSPSPGLSPFTSSSCFSFNPEE
205	1055	Λ.	2033	402	170	MKHYLHSQACSVFNYHLSPRTFPRYPGLMVP
<u> </u>						PLQCQMHPEESTQFSIKLQPPPVGRKNRERVE
						SSEESAP
284	1634	Α	2836	2	384	KTLPRTLLDILADGTILKVGVGCSEDASKLLQ
1 1						DYGLVVRGCLDLRYLAMRQRNNLLCNGLSL
1						KSLAETVLNFPLDKSLLLRCSNWDAETLTED
				İ		QVIYAARDAQISVALFLHLLGYPFSRNSPGEK
-						KR
285	1635	A	2843	20	271	PIRPYYSYSGLDRDCSWLPLAKAWLPDVMIL
1 1		l	1	ł		VCDRVSEDGINRQQAQEWCIKHGFELVELSP
286	1636	Ā	2845	197	270	EELPEEDGKCLCVRRKYGTYI
287	1637	A	2843	2	278 427	TAEDVLTVAYEHGVNLFDTAEVYAAGK
20'	1057	^	2021	-	461	FVAEVRREWAK YME VHEKASFTNSELHRAM NLHVGNLRLLSGPLDOVRAALPTPALSPKDK
.	ļ	j	J			AVLQNLKRILAKVQEMRDQRVSLEQQLRELI
] [ļ		QKDDITGSLVTTDHSQMKKLFEEQLKKYDQL
						KVYLEQNLAAQDRVLCALT
288	1638	A	2859	2	469	FVNLGILTCIECSGIHREMGAHISRIQSLELDK
	. }			ļ		LGTSELLPAKNVGNNSFNDIMEANLPSPSPKP
		ļ		ŀ	1	TPSSDMTVRKEYITAKYVDHRFSRKTCSTSSA
1 1	1	1	- 1	1	1	KLNELLEAIKSROLLALIQVYAEGVELMEPLL
						EPGQELAETALHLAVRTADQTSLHLVE \
289	1639	A	2861	2	454	FVASGGPATARMSDSQFFCVAEERSGHCAVV
	ł	-	- 1	ł		DGNFLYVWGGYVSIEDNEVYLPNDEIWTYDI
j ļ	ļ	ļ	}			DSGLWRMHLMEGELPASMSGSCGACINGKL
	j	1	- 1			YIFGGYDDKGYSNRLYFVNLRTRDETYIWEK
290	1640	Ā	2868	1	279	ITDFEGQPPTPRDKLSCWVYKDRLIYFG
250	1040	^	2000	• [378	FRQGQLYKVFLHGSQGQVYHSQQVGPPGSAI
[{		1		SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF PDCASCLQAQDPLCGWCVLQGRCTRKGQCG
L						I DOLLOCKA QUI DOG WOYL QURCTRA QUO

NO of No o	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
mucleotide seq							
			""				E-Dhanylalaning C-Chroine Hallistiding
		1					
Unified 1914							
maino acid residue of peptide requence maino acid residue of peptide sequence maino acid residue of peptide sequence maino acid residue of peptide sequence maino acid residue of peptide maino acid residue of pe		ucince		1		1	
Persidue of peptide sequence	dence	ļ		714			
peptide	ì				1		1=1 nreonine, v=valine, W=1 ryptophan,
		ĺ		1		sequence	
RAGQINQWLWSYEEDSHCLHIQSLLFGHHPR QE QE QE QE QE QE QE Q	1						
1641					sequence		
291	1	1	i			ł	
PFIPESIBSHPÖHPFUNKVENPECTENNCTS							
	291	1641	A	2870	1	385	FRYMPNNRQQLLRKRHIGNDIVTIVFQEPGAL
			i	l		l	PFTPKSIRSHFQHVFVIVKVHNPCTENVCYSV
1642 A 2877 3 188 RPTRPPPATTQSPESTMDTSLKKEKSAILDLYI PPPPAVPYSPR VA WHCHGML VSCWCHL 293 1643 A 2878 1 427 REKEEVEEEDK VKETEKEAEQEKEEDSL GAGTHPDAAIPSGERTCGSEGSRSVLDLVNYF LSPEKLTAENRYYCESCASLQDAEKVVELDS GEVYLLTLLESPEDLRTMRRKHLDDVSIPLL LRLPLAGGRQAYDL LRLPLAGGRQ	1	ł		i		l	GVSRSKDVPPFGPPIPKGVTFPKSAVFRDFLL
1642 A 2877 3 188 RPTRPPPATTQSPESTMDTSLKKEKSAILDLYI PPPPAVPYSPR VA WHCHGML VSCWCHL 293 1643 A 2878 1 427 REKEEVEEEDK VKETEKEAEQEKEEDSL GAGTHPDAAIPSGERTCGSEGSRSVLDLVNYF LSPEKLTAENRYYCESCASLQDAEKVVELDS GEVYLLTLLESPEDLRTMRRKHLDDVSIPLL LRLPLAGGRQAYDL LRLPLAGGRQ	i	ŀ	l			ŀ	AKVINAENAAHKSEKFRAMATRTRQEYLKD
PPPPAVPYSPRYVAVBCHGMLVSCWCHL]	1	1	1	
PPPPAYYSPRYVAWICHGMLVSCWCHEDS.	292	1642	Α	2877	3	188	RPTRPPPATTOSPESTMDTSLKKEKSAILDLYI
1643					ĺ		PPPPAVPYSPRYVAVHCHGMLVSCWCHI.
GAGTHPDAAIPSGERTCGSEGSRSVLDLVAVF	293	1643	Ā	2878	1	427	REKEEPVEEEDKVVKETEKEAEOEKEEDSI
LSPEKLTAENRYYCESCASI,QDAEKVVELSQ GPCYLLITLERSFDLRTMRRRKILDDVSIPLL LRLPLAGGRQAYDL LRLPLAGGRQAYDL LRLPLAGGRQAYDL LRLPLAGGRQAYDL LRLPLAGGRQAYDL LRLPLAGGRQAYDL GPCYLLITLERSFDLRTMRRKILDDVSIPLL LRLPLAGGRQAYDL LRLPLAGGRQAYDL GPCYLLITLERSFDLRTMRRKILDDVSIPLL LRLPLAGGRQAYDL LRLPLAGGRQAYDL LASSQHGILNNI: SLLFSICKTCIRTMDHHCPRA LASSQHGILNNI: SLLFSICKTCIRTMDHHCPRA LASSQHGILNNI: SLLFSICKTCIRTMDHHCPRA LASSQHGILNNI: SLLFSICKTCIRTMDHHCPRA LASSQHGILNNI: SLLFSICKTCIRTMDHHCPRA LASSQHGILNNI: SLLFSICKTCIRTMDHHCPRA LASSQHGILNNI: SLLFSICKTCIRTMDHHCPRA LASSQHGILNNI: SLLFSICKTCIRTMDHHCPRA LASSQHGILNNI: SLLFSICKTCIRTMDHHCPRA LASSQHGILNNI: SLLFSICKTCIRTMDHHCPRA LASSQHGILNNI: SLLFSICKTCIRTMDHHCPRA LASSQHGILNNI: SLLFSICKTCIRTMDHHCPRA LASSQHGILAGFILAGFILAGFILAGFILAGFILAGFILAGFILAG			1		•	1 '-'	
	1		ŀ	l			
LRIPLAGGROGAYDL						ļ	
1644				1			
1645	204	1644		2870	100	245	
295	294	1044	A	20/9	109	243	
NNCVGEQNHRFFCALHCKSKHFCIEFTLNTNF FNCFLPGAEKSTIDAPFSLQPFLQDSKYNTALS	205	17.45		2000			
PINCELPGAEKSTIDAPFSLQPFLQDSKYNTALS	293	1645	Α	2880	3	320	LASSQHGILNNLSLLFSICKTCIRTMDHHCPRA
LSESISQ	1 1]	
296]						
RIQEFSQKMDQVRGHWPVST							
297	296	1646	Α	2892	209	363	SQYSHSLDYHLLQVTKNPFTLGDSSNPGQTE
RLYSTMGRFLRDRRNPACREMAVVLLANLA QGDSLAARALVQKGISIGHLIGFLEDSLAAT QIQQSASLIHMINPPPEPISVDMMRRACRA LLALAKVDDNHSEF							RLQEFSQKMDQVRGHWPVST
	297	1647	Α	2893	8	424	SPXTLXLDTFILLGIQDNILVLILATPPFMAGG
QGDSLAARAIAVQKGSIGHLLGFLEDSLAAT QIQQSQASLLHMHNPPFEPTSVDMMRRACRA LLALAKVDDNHSEF	1 1				!		KLYSTMGRFLRDRKNPACREMAVVLLANLA
							OGDSLAARAIAVOKGSIGHLLGFLEDSLAAT
LLALAK V D D N H SEF	1				!		OIOOSOASLLHMHNPPFEPTSVDMMRRACRA
298	1 1						LLALAKVDDNHSEF
SGLLNASAQVNL	298	1648	A	2894	310	445	
1649							
GYFQAYNVLILTMQASLPKVLRFCACAGMIY	299	1649	A	2898	1	492	KIKAKNI TNYDI CSIFI GTSTI I VWVGVIRVI
LGYTFCGWIVLGPYHDKFENLNTVAECLFSL VNGDDMFATFAQIQQKSILVWLFSRLYLYSFI SLFIYMILSLFIALITDSYDTIKKFQQNGFPETD LQEF	i l				_		GYFOAYNVLILTMOASLPKVLRFCACAGMIV
VNGDDMFATFAQIQQKSILVWLFSRLYLYSFI SLFIYMILSLFIALITDSYDTIKKFQQNGFPETD LQEF							I GVTFCGWIVI GPVHDKEENI NTVAECI ESI
SLFIYMILSLFIALITDSYDTIKKFQQNGFPETD LQEF	!!	i					
LQEF							
1650	l l						
TVTVRFVNKADFPKVRAKEQTFMFPENQPVS SLVTTITGSSLRGEPMSYYIASGNLGNTFQIDQ LTGQVSISQPLDFEKIQKYVVWIEARDGGVPP FSSYEKLDITVLDVNDNAPIF	300	1650		2001		445	
SLVTTITGSSLRGEPMSYYIASGNLGNTFQIDQ	500	1050	n.	27UI	•	747	
LTGQVSISQPLDFEKIQKYVVWIEARDGGVPP FSSYEKLDITVLDVNDNAPIF 301 1651 A 2902 162 433 THFICLPLGYCFPLLDKDLQLPSGFNCNFDFLE EPCGWMYDHAKWLRTTWASSSSPNDRTFPG KPAVSEDMKELRPACSTYFNPRFPYKL 302 1652 A 2909 2 412 GPQMLCKKIYFIWYTRSQCGFEWLADIMQEV EENDHQDLVSVHIYVTQLAEKFDLRTTMLYI CERHFQKVLNRSLFTGLRSITHFGRPPFEPFN SLQEVHPQVRKIGVFSCGPPGMTKNVEKACQ LVNRQDRAHFM 303 1653 A 2914 291 453 KLNRWLCFFYSWSFGILLYEMVTLGAPPYPE VPPTSILEHLQRRKIMKRPSSCS 304 1654 A 2926 179 354 PGVPSQALRKAESLKKCLSVMEAKVKAQTAP NKDVQREIADLGEVGAASLPPSSGPGA 305 1655 A 2938 135 438 GMGYLHAKGILHKDLKSKNVFYDNGKVVIT DFGLFSISGVLQAGRREDKLRIQNGWLCHLA PEIIRQLSPDTEEDKLPFSKHSDVFALGTIWYE LHAREWP 306 1656 A 2944 2 329 VRWNSCVNCSCAFGNGASLSTSLGESSGCLW EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA	1	- 1			ľ	i	
1651							SLV1111GSSLRGEPMSYYIASGNLGNTFQIDQ
301 1651 A 2902 162 433 THFICLPLGYCFPLLDKDLQLPSGFNCNFDFLE EPCG WMYDHAK WLRTTWASSSSPNDRTFPG KPAVSEDMKELRPACSTYFNPRFPYKL 302 1652 A 2909 2 412 GPQMLCKKIYFIWVTRSQCQFEWLADIMQEV EENDHQDLVSVHIYVTQLAEKFDLRTTMLYI CERHFQK VLNRSLFTGLRSITHFGRPPFEPFN SLQEVHPQVRKIGVFSCGPPGMTKNVEKACQ LVNRQDRAHFM 303 1653 A 2914 291 453 KLNRWLCFFYSWSFGILLYEMVTLGAPPYPE VPPTSILEHLQRRKIMKRPSSCS 304 1654 A 2926 179 354 PGVPSQALRKAESLKKCLSVMEAKVKAQTAP NKDVQREIADLGEVGAASLPPSSGPGA 305 1655 A 2938 135 438 GMGYLHAKGILHKDLKSKNVFYDNGKVVIT DFGLFSISGVLQAGRREDKLRIQNGWLCHLA PEIIRQLSPDTEEDKLPFSKHSDVFALGTIWYE LHAREWP 306 1656 A 2944 2 329 VRWNSCVNCSCAFGNGASLSTSLGESSGCLW EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA		Ī			1		
BPCG WMYDHAK WLRTTWASSSSPNDRTFPG KPAVSEDMKELRPACSTYFNPRFPYKL 302	201	1651					
SPAVSEDMKELRPACSTYFNPRFPYKL SPAVSEDMKELRPACSTYFNPRFPYKL 302 1652 A 2909 2 412 GPQMLCKKIYFIWVTRSQCQFEWLADIMQEV EENDHQDLVSVHIYVTQLAEKFDLRTTMLYI CERHFQKVLNRSLFTGLRSITHFGRPPFEPFFN SLQEVHPQVRKIGVFSCGPPGMTKNVEKACQ LVNRQDRAHFM SVPTSILEHLQRRKIMKRPSSCS 304 1654 A 2914 291 453 KLNRWLCFFYSWSFGILLYEMVTLGAPPYPE VPPTSILEHLQRRKIMKRPSSCS 304 1654 A 2926 179 354 PGVPSQALRKAESLKKCLSVMEAKVKAQTAP NKDVQREIADLGEVGAASLPPSSGPGA 305 1655 A 2938 135 438 GMGYLHAKGILHKDLKSKNVFYDNGKVVIT DFGLFSISGVLQAGRREDKLRIQNGWLCHLA PEIIRQLSPDTEEDKLPFSKHSDVFALGTIWYE LHAREWP S06 1656 A 2944 2 329 VRWNSCVNCSCAFGNGASLSTSLGESSGCLW EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA	301	1001	A.	2902	162	433	
302 1652 A 2909 2 412 GPQMLCKKIYFIWVTRSQCQFEWLADIMQEV EENDHQDLVSVHIYVTQLAEKFDLRTTMLYI CERHFQKVLNRSLFTGLRSITHFGRPPFEPFFN SLQEVHPQVRKIGVFSCGPPGMTKNVEKACQ LVNRQDRAHFM 303 1653 A 2914 291 453 KLNRWLCFFYSWSFGILLYEMVTLGAPPYPE VPPTSILEHLQRRKIMKRPSSCS 304 1654 A 2926 179 354 PGVPSQALRKAESLKKCLSVMEAKVKAQTAP NKDVQREIADLGEVGAASLPPSSGPGA 305 1655 A 2938 135 438 GMGYLHAKGILHKDLKSKNVFYDNGKVVIT DFGLFSISGVLQAGRREDKLRIQNGWLCHLA PEIIRQLSPDTEEDKLPFSKHSDVFALGTIWYE LHAREWP 306 1656 A 2944 2 329 VRWNSCVNCSCAFGNGASLSTSLGESSGCLW EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA		į		j			
EENDHQDLVSVHIYVTQLAEKFDLRTTMLYI CERHFQKVLNRSLFTGLRSITHFGRPPFEPFFN SLQEVHPQVRKIGVFSCGPPGMTKNVEKACQ LVNRQDRAHFM SUMPLIFIED STATE OF STATE O	1						
BENDHQDLVSVHIYVTQLAEKFDLRTTMLYI CERHFQKVLNRSLFTGLRSITHFGRPPFEPFFN SLQEVHPQVRKIGVFSCGPPGMTKNVEKACQ LVNRQDRAHFM 303 1653 A 2914 291 453 KLNRWLCFFYSWSFGILLYEMVTLGAPPYPE VPPTSILEHLQRRKIMKRPSSCS 304 1654 A 2926 179 354 PGVPSQALRKAESLKKCLSVMEAKVKAQTAP NKDVQREIADLGEVGAASLPPSSGPGA 305 1655 A 2938 135 438 GMGYLHAKGILHKDLKSKNVFYDNGKVVIT DFGLFSISGVLQAGRREDKLRIQNGWLCHLA PEIIRQLSPDTEEDKLPFSKHSDVFALGTIWYE LHAREWP 306 1656 A 2944 2 329 VRWNSCVNCSCAFGNGASLSTSLGESSGCLW EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA	302	1652	A	2909	2	412	GPQMLCKKIYFIWVTRSQCQFEWLADIMQEV
SLQEVHPQVRKIGVFSCGPPGMTKNVEKACQ LVNRQDRAHFM 303 1653 A 2914 291 453 KLNRWLCFFYSWSFGILLYEMVTLGAPPYPE VPPTSILEHLQRRKIMKRPSSCS 304 1654 A 2926 179 354 PGVPSQALRKAESLKKCLSVMEAKVKAQTAP NKDVQREIADLGEVGAASLPPSSGPGA 305 1655 A 2938 135 438 GMGYLHAKGILHKDLKSKNVFYDNGKVVIT DFGLFSISGVLQAGRREDKLRIQNGWLCHLA PEIIRQLSPDTEEDKLPFSKHSDVFALGTIWYE LHAREWP 306 1656 A 2944 2 329 VRWNSCVNCSCAFGNGASLSTSLGESSGCLW EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA		l					EENDHQDLVSVHIYVTQLAEKFDLRTTMLYI
SLQEVHPQVRKIGVFSCGPPGMTKNVEKACQ LVNRQDRAHFM 303 1653 A 2914 291 453 KLNRWLCFFYSWSFGILLYEMVTLGAPPYPE VPPTSILEHLQRRKIMKRPSSCS 304 1654 A 2926 179 354 PGVPSQALRKAESLKKCLSVMEAKVKAQTAP NKDVQREIADLGEVGAASLPPSSGPGA 305 1655 A 2938 135 438 GMGYLHAKGILHKDLKSKNVFYDNGKVVIT DFGLFSISGVLQAGRREDKLRIQNGWLCHLA PEIIRQLSPDTEEDKLPFSKHSDVFALGTIWYE LHAREWP 306 1656 A 2944 2 329 VRWNSCVNCSCAFGNGASLSTSLGESSGCLW EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA	j		ļ	J		J	CERHFQKVLNRSLFTGLRSITHFGRPPFEPFFN
LVNRQDRAHFM 303 1653 A 2914 291 453 KLNRWLCFFYSWSFGILLYEMVTLGAPPYPE VPPTSILEHLQRRKIMKRPSSCS 304 1654 A 2926 179 354 PGVPSQALRKAESLKKCLSVMEAKVKAQTAP NKDVQREIADLGEVGAASLPPSSGPGA 305 1655 A 2938 135 438 GMGYLHAKGILHKDLKSKNVFYDNGKVVIT DFGLFSISGVLQAGRREDKLRIQNGWLCHLA PEIIRQLSPDTEEDKLPFSKHSDVFALGTIWYE LHAREWP 306 1656 A 2944 2 329 VRWNSCVNCSCAFGNGASLSTSLGESSGCLW EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA			:	1			SLOEVHPOVRKIGVFSCGPPGMTKNVEKACO
1653 A 2914 291 453 KLNRWLCFFYSWSFGILLYEMVTLGAPPYPE VPPTSILEHLQRRKIMKRPSSCS 304							
VPPTSILEHLQRRKIMKRPSSCS 304 1654 A 2926 179 354 PGVPSQALRKAESLKKCLSVMEAKVKAQTAP NKDVQREIADLGEVGAASLPPSSGPGA 305 1655 A 2938 135 438 GMGYLHAKGILHKDLKSKNVFYDNGKVVIT DFGLFSISGVLQAGRREDKLRIQNGWLCHLA PEIIRQLSPDTEEDKLPFSKHSDVFALGTIWYE LHAREWP 306 1656 A 2944 2 329 VRWNSCVNCSCAFGNGASLSTSLGESSGCLW EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA	303	1653	Α	2914	291	453	
304 1654 A 2926 179 354 PGVPSQALRKAESLKKCLSVMEAKVKAQTAP NKDVQREIADLGEVGAASLPPSSGPGA 305 1655 A 2938 135 438 GMGYLHAKGILHKDLKSKNVFYDNGKVVIT DFGLFSISGVLQAGRREDKLRIQNGWLCHLA PEIIRQLSPDTEEDKLPFSKHSDVFALGTIWYE LHAREWP 306 1656 A 2944 2 329 VRWNSCVNCSCAFGNGASLSTSLGESSGCLW EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA		ŀ					
NKDVQREIADLGEVGAASLPPSSGPGA NKDVQREIADLGEVGAASLPPSSGPGA SOS 1655 A 2938 135 438 GMGYLHAKGILHKDLKSKNVFYDNGKVVIT DFGLFSISGVLQAGREEDKLRIQNGWLCHLA PEIIRQLSPDTEEDKLPFSKHSDVFALGTIWYE LHAREWP 306 1656 A 2944 2 329 VRWNSCVNCSCAFGNGASLSTSLGESSGCLW EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA	304	1654	A	2926	179	354	
305 1655 A 2938 135 438 GMGYLHAKGILHKDLKSKNVFYDNGKVVIT DFGLFSISGVLQAGRREDKLRIQNGWLCHLA PEIIRQLSPDTEEDKLPFSKHSDVFALGTIWYE LHAREWP 306 1656 A 2944 2 329 VRWNSCVNCSCAFGNGASLSTSLGESSGCLW EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA					• • •	JJ7	
DFGLFSISGVLQAGRREDKLRIQNGWLCHLA PEIIRQLSPDTEEDKLPFSKHSDVFALGTIWYE LHAREWP 306	305	1655	<u> </u>	2028	125	A20	
PEIIRQLSPDTEEDKLPFSKHSDVFALGTIWYE LHAREWP 306	505	1000	^	2730	133	430	
306 1656 A 2944 2 329 VRWNSCVNCSCAFGNGASLSTSLGESSGCLW EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA		ļ		l		1	
306 1656 A 2944 2 329 VRWNSCVNCSCAFGNGASLSTSLGESSGCLW EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA		. }	l	ŀ	1	i	
EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA	206	1000		2014			
	300	1000	A	2944	2	329	
LTKGALWAVFLLAGSALLCAEVTGVIWRQPE	1	1	ŀ		Į.	j	
							LTKGALWAVFLLAGSALLCAEVTGVIWRQPE

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	nou .	in in	nucleotide	location	
	1					F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ł	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1	i	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	1	1	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	1	1	ļ	peptide	•	/=possible nucleotide deletion, \=possible
}	J		i	sequence		nucleotide insertion
	 		 		 	SKTKLSFKVSSSA
307	1657	A	2950	2	411	NYLCIAKNSAGSAMGKTRLVVQVPPVIENGL
1 307	1057	Α	2930	4	411	
1						PDLSTTEGSHAFLPCKARGSPEPNITWDKDGQ
1	{	1	Í	ĺ	i	PVSGAEGKFTIQPSGELLVKNLEGQDAGTYT
		i	İ			CTAENAVGRARRRVHLTILVLPVFTTLPGDRS
	<u></u>	<u> </u>	1		1	LRLGDRLWLR
308	1658	Α	2951	1	407	PTRPPRVRFDNEFDAESQRKRTTSVSKMERM
	İ	1	1			DSSLPEEEEDEDKEAINGSGNAENRERHSESS
			ļ			DWMKTVPSYNQTNSSMDFRNYMMRDETLEP
	ļ	j	ļ		j	LPKNWEMAYTDTGMIYFIDHNTKTTTWLDP
			İ			
309	1659	 	2954		150	RLCKKAKAPEDC
309	1058	Α	2954	2	179	QDFLTLTLTEPTGLLYVGAREALFAFSMEALE
L	<u> </u>	<u> </u>				LQGAVRGGAVGGSRACQRARPRGAVLG
310	1660	A	2959	1	419	QDMMERAIIDTFVGHDVVEPGSYVQMFPYPC
1	1	1	1		[YTRDDFLFVIEHMMPLCMVISWVYSVAMTIO
			l		Ì	HIVAEKEHRLKEVMKTMGLNNAVHWVAWFI
						TGFVQLSISVTALTAILKYGQVLMHSHVVIIW
1	1				ł	LFLAVYAVATIMFCF
311	1661	A	2963	3	465	
1 211	1001	, A	2903	3	403	MKPQMPGLGAPNGYGPGRGRAGVPGGPERR
1	1	}				PWVPHLLPFSSPGYLGVMKAQKPGAGEGMK
ŀ					'	PQKPGLRGTLKPQKSGHGHENGPWPGPCNA
1 .						RVAPMLLPRLPTPGVPSDKEGGWGLKSQPPS
1	}] .		Ì	AVQNGKLPGHQPPNGYGPGAEPGFNGGLEPQ
L	1				•	KI .
312	1662	Α	2967	3	405	WLAQEWSPCTVTCGQGLRYRVVLCIDHRGM
					,	HTGGCSPKTKPHIKEECIVPTPCYKPKEKLPV
1	}		1 1			EAKLPWFKQAQELEEGAAVSEEPSFIPEAWS
l	ì					ACTVTCGVGTQVRIVRCQVLLSFSQSVADLPI
313	1663	A	2969	2	420	DECEGPKPA
313	1003	A	2909	2	430	VVADNCROGYLDALRFLERRGLTKEPVLWT
 						LVSKEPPAPADGNWDAGCDQRRKGGLSLNW
}						KVPHVQVKDVPNFEQLSPELEAALKKACTRD
			[[PSRWARFWHSGPGQVLTYLLLPCTLPFEYIYF
L						RSRRLVVWLPDVPADLWWMQ
314	1664	Α	2971	422	33	LDXSHNALQRLRPGWLAPLFQLRALHLDHNE
1 1	,	'				LDALGRGVFVNASGLRLLDLSSNTLRALGRH
						DLDGLGALEKLLLFNNRLVHLDEHAFHGLRA
1			1			LSHLYLGCNELASFSFDHLHGLSATHLLTLDL
				l		COVID 4
315	1665	_	2072	 	505	SSNRM
1 212	1003	A	2973	1	525	ITVSTHASGSPFGLEPQSGWLWVRAALDREA
1			!			QELYILKVMAVSGSKAELGQQTGTATVRVSI
1					l	LNQNEHSPRLSEDPTFLAVAENQPPGTSVGRV
j i						FATDRDSGPNGRLTYSLQQLSEDSKAFRIHPQ
[[ĺ	ſ	TGEVTTLQTLDREQQSSYQLLVQVQDGGSPP
						RSTTGTVHVAVLDLNDNT
316	1666	A	2978	2	400	ELVVELVSAGKSGPERNTYEVQVVTGNVPKA
1 1		••	2770	-	₹ 700 ,	CTDANIAN TRACEPACRACERAL RESPONSE
i	•					GTDANVYLTIYGEEYGDTGERPLKKSDKSNK
	-			,		FEQGQTDTFTIYAIDLGALTKIRIRHDNTGNR
)				j		AGWFLDRIDITDMNNEITYYFPCQRWLAVEE
]	DDGQLSRE
317	1667	A	2981	3	440	VLNCQGRPTRPVRINGDGQEVLYLAESDNVR
i						LGCPYVLDPDDYGPNGLDIEWMQVNSNPAH
[[ĺ	1	1	į	HRENVFLSYQDKRINHGSLPHLQHRVRFAAS
[1				l		
]			ľ	ŀ	ĺ	DPSQYDASINLMNLQVSDTATYECRVKKTTM
	166	_	2005			ATRKVIVTVQARPAVPMCWTEGQ
318	1668	Α	2995	119	414	LPEKEFPIIRKSSSLKVTKCLFTEQPKPIIILRFA
		ļ	j		ļ	ENYDARLLRIDIANTLREQVQELFNKTYGKQ
					i	RRTPGEGHVAAVDREVAGFPVPAEGISGETIH
319	1669	A	2999	2	332	GFFAYTYGRLVVVEDLHSGAQQHWSGHSAEI

SEQ ID	SEQ ID	Met	SEQ	Predicted	I W 4" - 1 - 1	
NO: of	NO: of	hod	ID NO:	beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid. E=Glutamic Acid.
nucl-	peptide	1 1100	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1		l		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		į .		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide]	/=possible nucleotide deletion, \=possible
	ļ			sequence	<u> </u>	nucleotide insertion
	1	ł	1	1	l	STLALSHSAQVLASASGRSSTTAHCQIRVWD
	1	ļ				VSGGLCQHLIFPHSTTVLALAFSPDDRLLVTL
320	1670	Ā	3000	693	322	GDHDGRTLALWGTGHL
320	1070	1	3000	093	322	IDESTGLIITVNYLDYETKTSYMMNVSATDQA PPFNQGFCSVYITLLNELDEAVQFSNASYEAA
}	1	ł		}		ILENLALGTEIVRVQAYSIDNLNQITYRFDAY
			-			TSTQAKALFKIDAITVRGWGQGAPFFPI
321	1671	A	3001	6	383	RIPRGKACXTVLGRSTGELEGFASSRLPPOPC
1			1			GWGQSSDLLSRIDLDELMKKDEPPLDFPDTLE
1			1	1	ŀ	GFEYAFNEKGQLRHIKTGEPFVFNYREHLHR
				_	İ	WNQKRYEALGEIITKYVYELLEKDCNSKKVS
322	1672	Α	3007	192	447	ERVRNSLFPGRGDSQCACCPSSPVWVFLETGF
						LFPWLFLQVEVIKKAYMQGEVEFEDGENGK
	<u> </u>					DGAASPRNVGHNIYILAHQLARH
323	1673	Α	3019	18	245	KELLFYHLIVNNINFFNTRYAKIHIPIIASVSEH
	·	ĺ				QPTTWVSFFFDLHILVCTFPAGLWFCIKNIND
324	1674	A	3020	500	707	ERVFGKRGF
324	10/4	A	3020	523	797 .	LCYFSARYHQRKIFGILYIFTLSAINRKEPNLFI
			1			YLFIFFEMESHSVTHAGVQRHNLNSLQPLPPG FKRFSCLCFLSSWNYRGAPPGPANF
325	1675	A	3022	2	156	NDFLPLYFGWVLTKKSSETLRKAGQVFLEEL
323	1073	^	3022	-	150	GNHKAFKKELRQCRWQVGAL
326	1676	A	3023	38	172	KMVRGSKKLISFFPGGPYGILAGRDPSKGLAT
		``	***	20	172	FCLNKEALKDEFE
327	1677	A	3027	1	385	LTLEFLLLPAASELAHGKRLACCIVDHKLPEC
						GFYGLYDKILLFKHDPTSANLLQLVRSSGDIQ
		j	ļ ·			EGDLVEVVLSASATFEDFQIRPHALTVHSYRA
						PAFCDHCGEMLFGLVRQGLKCDGCGLNYHK
328	1678		2020			RC
320	10/8	Α	3030	13	569	ITRPTISCQRPGPGLAAGMLPYTVNFKVSART
						LTGALNAHNKAAVDWGWQGLIAYGCHSLV
						VVIDSITAQTLQVLEKHKADVVKVKWAREN YHHNIGSPYCLRLASADVNGKIIVWDVAAGV
						AQCEIQEHAKPIQDVQWLWNQDASRDLLLAI
						HPPNYIVLWNADTGTKLWKKSYADNILSFSF
						D
329	1679	Α	3038	90	744	SVNLPPSLWPWEEAMDSTKSEPLKGSPEAED
						GNIEYKKLVNPSQYRFEHLVTQMKWRLQEG
				.		RGEAVYQIGVEDNGLLVGLAEEEMRASLKTL
}					ļ	HRMAEKVGADITVLREREVDYDSDMPRKITE
					ĺ	VLVRKVPDNQQFLDLRVAVLGNVDSGKSTL
				i		LGVLTQGELDNGRGRARLNLFRHLHEIQSGR
						TSSISFEILGFNSKGEVHGINGTQWGQTLRMG
330	1680	Α	3040	3	397	L CCTLLLLTINGWAY COTTLATECCRILLANDERS
330	1000	^	DFVC	3	371	LCSTLLLLTIPSWVLSQITLKESGPTLMKPTET LTLTCTFSGFSLNTSGVGVAWIRQPPGKALE
			1			WLALIYWDDDKRYSPSLNDRLTIAKDTSRNO
		-	l	ľ		VVLTMTNMGPVDTATYYCAQFARGARGSN
J		1			ł	WFDPWGQ
331	1681	A	3043	3	1509	AGIRHEAPPTTSNRHRRQIDRGVTHLNISGLK
		[İ		MPRGIAIDWVAGNVYWTDSGRDVIEVAQMK
						GENRKTLISGMIDEPHAIVVDPLRGTMYWSD
	1	l				WGNHPKIETAAMDGTLRETLVQDNIQWPTG
1	1	ſ		ĺ		LAVDYHNERLYWADAKLSVIGSIRLNGTDPI
1]			1	VAADSKRGLSHPFSIDVFEDYIYGVTYINNRV
1		1	j	ļ		FKIHKFGHSPLVNLTGGLSHASDVVLYHQHK
1				1		QPEVTNPCDRKKCEWLCLLSPSGPVCTCPNG
1		ı		1	1	KRLDNGTCVPVPSPTPPPDAPRPGTCNLQCFN
				l		GGSCFLNARRQPKCRCQPRYTGDKCELDQC

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion WEHCRNGGTCAASPSGMPTCRCPTGFTGPKC TQQVCAGYCANNSTCTVNQGNQPQCRCLPG
1						FLGDRCQYRQCSGYCENFGTCQMAADGSRQ CRCTAYFEGSRCEVNKCSRCLEGACVVNKQS GDVTCNCTDGRVAPSCLTCVGHCSNGGSCT MNSKMMPECQCPPHMTGPRCEEHVFSQQQP GHIASILIP
332	1682	A	3045	3	952	TTTISNFHTQVNRTYCCGTYRAGPMRQISLVG AVDEEVGDYFPEFLDMLEESPFLKMTLPWGT LSSLRLQCRSQSDDGPIMWVRPGEQMIPTAD MPKSPFKRRRSMNEIKNLQYLPRTSEPREVLF EDRTRAHADHVGQFDWQSTAAVGVLKAV QFGEWSDQPRITKDVICFHAEDFTDVVQRLQ LDLHEPPVSQCVQWVDEAKLNQMRREGIRY ARIQLCDNDIYFIPRNVIHQFKTVSAVCSLAW HIRLKQYHPVVEATQNTESNSNMDCGLTGKR ELEVDSQCVRIKTESEEACTEIQLLTTASSSFP PASE
333	1683	A	3046	497	167	SACSTGPELPGRATRSLTRPANQKGCDGDRL YYDGCAMIAMNGSVFAQGSQFSLDDVEVLT ATLDLEDVRSYRAEISSRNLAVSAPVDTCVG CSSKTWKVAPFVRAWWRP
334	1684	A	3053	37	276	VITDLEEQLNQLTEDNAELNNQNFYLSKQLD EASGANDEIVQLRSEVDHLRREITEREMQLTS QKQVRRVNKVVRSLEDF
335	1685	A	3054	2	846	WDAWGDWSDCSRTCGGGASYSLRRCLTGR NCEGQNIRYKTCSNHDCPPDAEDFRAQQCSA YNDVQYQGHYYEWLPRYNDPAAPCALKCH AQGQNLVVELAPKVLDGTRCNTDSLDMCISG ICQAVGCDRQLGSNAKEDNCGVCAGDGSTC RLVRGQSKSHVSPEKREENVIAVPLGSRSVRI
						TVKGPAHLFIESKTLQGSKGEHSFNSPGVFVV ENTTVEFQRGSERQTFKIPGPLMADFIFKTRY TAAKDSVVQFFFYQPISHQWRQTDFFPCTVT CGGG
336	1686	A	3058	54	347	VVGKQEAGAHSDSCCLLHTPPRLTPAHSRKA LRNSRIVSQKDDVHVCIMCLRAIMNYQVSRG AWDWRLGSPACPHWGLHKLPRLWDPLSLYP VLCWGT
337	1687	A	3059	2	709	ILTSLVELTRFETLTPRFSATVPPCWVEVQQE QQQRRHPQHLHQQHHGDAAQHTRTWKLQT DSNSWDEHVFELVLPKACMVGHVDFKFVLN SNITNIPQIQVTLLKNKAPGLGKVNGLRLCPF LEDHKEDILCGPVWLASGLDLSGHAGMLTLT SPKLVKGMAGGKYRSFLIHVKAVNERGTEEI CNGGMRPVVRLPSLKHQSNKGYSLASLLAK VAAGKEKSSNVKNENTSGTRK
338	1688	A	3060	85	384	KAFYNYHVLELLQMLVTGGVSSQLEQHLDK DKVYGVADSCTSLLSGRNRCKLGLLSLHETIL SDVNPRNTFGQLFCGSLDLFGILCVGLYRIIDE EELNP
339	1689,	A	3063	236	362	CFLCLSGDFMVMTIFFNVSRRFGYVAFQNYV PSSVTTMLSWV
340	1690		3065	3	1249	DLWQFTPLHEAASKNRVEVCSLLLSYGADPT LLNCHNKSAIDLAPTPQLKERLAYEFKGHSLL QAAREADVTRIKKHLSLEMVNFKHPQTHETA LHCAAASPYPKRKQICELLLRKGANINEKTKE FLTPLHVASEKAHNDVVEVVVKHEAKVNAL DNLGQTSLHRAAYCGHLQTCRLLLSYGCDPN

	EQ Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of NO: of hod II nucl- peptide in	O NO: beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
	nucleotide SSN location	location corresponding	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine.
	9/496 correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
1 ' 1 1 1	14 ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	peptide		/=possible nucleotide deletion, \=possible
	sequence		nucleotide insertion
			IISLQGFTALQMGNENVQQLLQEGISLGNSEA
			DRQLLEAAKAGDVETVKKLCTVQSVNCRDIE
		1	GRQSTPLHFAAGYNRVSVVEYLLQHGADVH
			AKDKGGLVPLHNACSYGHYEVAELLVKHGA VVNVADLWKFTPLHEAAAKGKYEICKLLLQ
			HGADPTKKNRDGNTPLDLVKDGDTDIQDLLR
			GDAALLDAAKKGCLARVKKLSSPDNVNCRD
		1	TQGRHSTPLHLAGK
341 1691 A 30	070 1	547	GVLIPSFQNQLFADILAGIESVTSEHNYQTLIA
1 1 1 1			NYNYDRDSEEESVINLLSYNIDGIILSEKYHTI
			RTVKFLRSATIPVVELMDVQGERLDMEVGFD
		ļ	NRQAAFDMVCIMLEKRVRHKILYLGSKDDT
1 1 1	ł		RDEQRYQGYCDAMMLHNLSPLRMNPRAISSI
342 1692 A 30	073 463	3	HLRMQLMRDALSANPDLDGVFCTN
342 1092 A 30	1/3 403	3	RINRCRKPSDADILVPGDTISLIGTTSLRIDYNE
			IDDNRVTAEEVDILLREGEKLAPVMAKTRILR AYSGVRPLVASDDDPSGRNVSRGIVLLDHAE
]	RDGLDGFITITGGKLMTYRLMAEWATDAVC
			RKLGNTRPCTTADLALPGSQEPAKVP
343 1693 A 30	75 250	1	LLIYLAIFAPVAMSALAGVKSVQQVRIRAAQS
			LGASRAQVLWFVILPGALPEILTGLRIGLGVG
		<u> </u>	WSTLVAAELIAATRGLGFM
344 1694 A 30	076 2	138	LYFDAYLQSLQVAAISTFCCLLIGYPLAWAV
			AHSKPSTRNILLLL
345 1695 A 30	078 469	3	LKIRGQRIELGEIDRVMQALPDVEQAVTHAC
			VINQAAATGGDARQLVGYLVSQSGLPLDTSA
			LQAQLRETLPPHMVPVVLLQLPQLPLIANGKL DRKALPLPELKAQAPGRAPKAGSETIIAAAFS
1 1 1			SLLGCDVQDADADFFALGGHSLLAMKLAT
346 1696 A 30	82 404	2	QNITSKDLDVRLDPQTVPIELEQLVLSFNHMI
			ERIEDVFTRQSNFSADIAHEIRTPITNLITQTEI
			ALSQSRSQKELEDVLYSNLEELTRMAKMVSD
			MLFLAQADNNQLIPEKKMLNLAHEVGKVFD
			QFEALPE
347 1697 A 30	84 3	340	NELTFKEAEISKLYTKVHPAYRTLLEKRQALE
			DEKAKLNGRVTAMPKTQQEIVRLTRDVESGQ
!!!!			QVYMQLLNKEQELKITEASTVGDVRIVDPAIT
348 1698 A 30	86 723	10	QPGVLKPKKGLIILGAI
346 1096 A 30	123	1	TQAMVWQQKACAEDDPQLSGRHWLHAATL YNIAAYPHLKGDDLAEQAQALSNRAYEEAA
			QRLPGTMRQMEFTVPGGAPITGFLHMPKGDG
			PFPTVLMCGGLDAMQTDYYSLYERYFAPRGI
			AMLTIDMPSVGFSSKWKLTQDSSLLHQHVLK
]]]			ALPNVPWVDHTRVAAFGFRFGANVAVRLAY
] ,			LESPRLKAVACLGPVVHTLLSGLKCQQQVPE
340 1500	07 0	040	MYLDVLASRLGMHDASTKSSTRENH
349 1699 A 30	87 2	249	RIRSSDPEITLAGTPLHAAYLIGMTLICAGFSV
	ĺ		GFGVAMSQALGPFSLRAGVASSTLGIAQVCG
350 1700 A 30	99 3	424	SSLWIWLAAVVGIGAWNM EAPEATPOPSOPGPSSPISLSAEEENAEGEVSR
1,00 A 30	´´ '	727	ANTPDSDITEKTEDSSVPETPDNERKASISYFK
			NQRGIQYIDLSSDSEDVVSPNCSNTVOEKTFN
			KDTVIIVSEPSEDEESQGLPTMARRNDDISELE
			DLSGMEDLK
351 1701 A 316	08 2	404	IKKNHIIGYQLLHRRALFEKRTRLSDYALIFG
		ſ	MFGIVVMVIETELSWGAYYKAPLYSLALKCL
			ISLFTIILLGLTIVYHAREIQLFMANYGADDWR
, , l	1	1	SALTYEPIFLILLEALRGVIHATPCRVSLSLWD
1 1 1	}	j	GLDLP

COPO ID	LODOTO	157	1 000	T 8 10	·	
SEQ ID NO: of	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	l	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
цепсе			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
l	1	1	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
ŀ	1	1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
İ	1	1		peptide		/=possible nucleotide deletion, \=possible
	1500	<u> </u>	ļ	sequence		nucleotide insertion
352	1702	A	3110	341	2	AQLAEVCPPQTLLTTNTSSISITAIAAEIKNPER
		1				VAGLHFFNPAPVMKLVEVVSGLATAAEVVE
1	1	ļ	}	}]	QLCELTLSWGKQPVRCHSTPGFIVNRVARPY
	<u> </u>		<u> </u>			YSEAWRALEEQVAAPEVI
353	1703	A	3111	3	188	HFSLFRIAFAVFLTYMTVGLPLPVIPLFVHHEL
	<u> </u>	ļ.,	<u> </u>	ļ	}	GYGNTMVGIAVGIQFLATVLTRGYAGRLA
354	1704	A	3116	367	225	WQLFHLNGTFLNIGETDTESCVNGWVYDRSS
L	1	1	1	1	ł	FPFSNMTEVRGLVFLS
355	1705	Α	3117	101	53	VINLVYLISSPRPELKPVDKESEVVMKFPDGF
1			1		1	EKFSPPILQLDEVDFYYDPKHVIFSRLSVSADL
	}	ł	l		Ì	ESRICVVGENGAGKSTMLKLLLGDL\APVRGI
ı						RHAHRNLKIGYFSQHHVGAAGT*TFSACGNL
1			1			LGTQVFLGRPEEEY\RHQLGFGMGISGELGHA
i i	1	}		ł		SSLPACLGGQKEAEVAFCSDGLLPCPNFL\IL\
						DEPTN\HLGHGRAIEALGPCLQTISGVGVILVS
			ı			HE*SALSRLVCRE\LWVC*GRSTSPF
356	1706	A	3121	137	466	RGGRDWGEHNQRLEEHQARAWQGAMDAG
1	1,00	1 **	3121	13,	1 400	AASREHARWQGTGLAPGTRVAVAPTCVOGL
}	}	J	1		J	PQERSVCRPFFSSRWREGPVWALGAGAHGKP
i						RWSGGVRCVVRGGRWFTPAPH
357	1707	A	3124	1249	229	
337	1707	Α	3124	1249	229	MLEAPGPSDGCELSNPSASRVSCAGQMLEVQ
1	1	ĺ	1 1	ĺ		PGLYFGGAAAVAEPDHLREAGITAVLTVDSE
	1		1			EPSFKAGPGVEDLWRLFVPALDKPETDLLSH
1			1 1			LDRCVAFIGQARAEGRAVLVHCHAGVSRSV
1	ì	l	1 1	1		AIITAFLMKTDQLPFEKAYEKLQILKPEAKMN
1		1	1			EGFEWQLKLYQAMGYEVDTSSAIYKQYRLQ
İ	1	l	1			KVTEKYPELQNLPQELFAVDPTTVSQGLKDE
1	l l	}	1 1			VLYKCRKCRRSLFRSSSILDHREGSGPIAFAH
į	1.		1			KRMTPSSMLTTGRQAQCTSYFIEPVQWMESA
Í	1	ĺ	1			LLGVMDGQLLCPKCSAKLGSFNWYGEQCSC
360	1700			-		GRWITPAFQIHKNRVDEMKILPVLGSQTGKI
358	1708	A	3127	816	139 .	EVETLGPRTPGP/EAQSPTPGSCPGWQEPSPGP
1	} :	l	ł			TPPP*LSGPGPQGAPVLGKLLPDPEETPAGKTP
1						LGKHFWWGL\PVTSANFSPGAAA*FGGALSPP
	į '		1			GGDL/GHMLLQGPPSPFRLQQQ*QTPPGSHSP
	1	l	1 1			PTANREINPGPAAAADTRSCWGHKRSWRGW
	j		i i			RGLAPWRLGFGSPGIP*PAPAGIP/GRPTWEGG
					•	KGAGGKPSETLTRSPPVWRGKRGSANGFLSW
	<u> </u>					VQILQ
359	1709	A	3132	3	191	HEHLLLLLCVFLVKSQGVNDNEEGFFSARG
1				ļ		HRPLDKKREDAPNLRPALADUTVCDYRAQIA
		L		ļ		*AASTPKRAASIAHNAVSCR*AQIA
360	1710	A	3134	1	286	REPPRPALLFF*DRVSLCCPGWNAVVOSOLT
) .	İ]]	ſ	i	AAPTSQVQ/SDSPTFPSSWDYRHVPEYPANFL
	'					*RQGFPMLPRLVSNSWAQTVHPPRPPKVLDL
					i	QA
361	1711	A	3135	56	1449	PVPAPRVSPSARGAPGRPRLPGVRGPRHS/WA
						AD*RGSRM/PPRAPAPSPTGP/APGGKKVRGR
	1 1		1 1	1		VPEDPDAYEPRCSAL*V*PTHVTSPOFCDP*N
			!!			GQIRSYFTVLLRGLNETMLVK/PLCRREP/PEA
	[[- 1	Í	GPGRQSTPAVTRDHRQHEDPRGAGRQWDAD
1			1			PRPSAP/PAEVATGSRPGRHMWMRLCLAAOO
						APGLPHRTSIRPGWRRLTEPEAWARRHRRPW
1	1 1			1	1	
	. !		ı i		ļ	GQRGAVRPPPQGAAPPPSHQGRRTNTDPSAT
	1		'			
<u> </u>				1	1	PRLTVMSRCLAPDLKAPASGPRGWRRGMPQ
						SS/GALLWTPPPTPRGSHSPRPREAPLRAIHPA
			,			SS/GALLWTPPPTPRGSHSPRPREAPLRAIHPA GPSK/SRAGASGRLPEVIYGWVTLFTPPEAGT
						SS/GALLWTPPPTPRGSHSPRPREAPLRAIHPA

NO: of No: of nucle-cotide contide of the contine o	CCCAID	SEQ ID	I Mat	CEO	D., 31-4-3	I D. C. J.	
nucleotide seq- uence USSN 09496 1914 uence USSN 09496 1914 uence USSN 09496 1914 uence USSN 09496 1914 uence USSN 09496 1914 uence USSN 09496 1914 uence USSN 09496 1914 uence USSN 09496 1914 uence USSN 09496 1914 uence USSN 09496 1914 uence USSN 09496 1914 uence USSN 09496 1914 uence USSN 09496 1914 uence USSN 09496 1914 uence USSN 09496 1914 USSN 09496 USSN 09496 uence USSN 09496 USSN 09496 USSSR 09406 USSSR USSSR 09406 USSSR USSSR 09406 USSSR USSSR 09406 USSSR USSSR USSR USSSR USSR USS	SEQ ID		Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Sequence		1	l nou	I		ľ	
Sequence		1	1				1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1914 ng no first araino acid residue of peptide pe			l			1 0	
amino acid residue of peptide residue of peptide residue of peptide sequence T-T-Thronine, V-Valine, W-Typtophan, Y-Typtophan, Y-Stop codon, peptide sequence Petide sequence Petide seq	1 -	delice	1				
residue of peptide pep	uence		1	914			
peptide	i						1=1 hreonine, V=Valine, W=1 ryptophan,
	į					sequence	
RNGSPLGNDTLSSGLPMGPRRQVWPLARVG							
GRISSPREPOVLKKELWQQTDIAGVGSASLYP DNL DNL		ļ. <u></u>	Ļ	ļ	sequence		
1712 A 3136 1270 274 RVGMVLGTREVGDSTPPPSPPLYPFTGNEFVQ HNTWQLSRVYPSDLRTDSSNYNPQELWNAG CQMV*GGSRWEGVEQVQONKPSSDGR VGECSRKLLG*EMLSVDITSRYRAPSTYLLNS LKEGLEGLHGSSCSSFLLG*EWAMNQTAGL EMDICDOHPRONGGGVVLKPDIELTDIGSSF HPEKPISPFKAGTLLNQVISVQJPKVDKTKE GSVDPLVKVQJGFQVLLATAQGTSTYVENDG FNPYWGGTLC*FRVLGPDFPM.FFGKMDYDW KSRNDLGKTPC*PGTCMQQOYNHIHLSKDG SILR*PASIFVYICQEGLEGDES APAMESSPLAPH-HQHLV*SVFQVLTLLIGV*** A 3140 57 418 MPAGSYGKSMFSSKKVINCLPKWRYHFVIA PAMNESPLAPH-HQHLV*SVFQVLTLLIGV** A MPAGSYGKSMFSSKKVINCLPKWRYHFVIA PAMNESPLAPH-HQHLV*SVFQVLTLLIGV** A MFAGSYGKSMFSSKKVINCLPKWRYHFVIA PAMNESPLAPH-HQHLV*SVFQVLTLLIGV** A MFAGSYGKSMFSSKKVINCLPKWRYHFVIA A MFAGSYGKSMFSSKKVINCLPKWRYHFVIA PAMNESPLAPH-HQHLV*SVFQVLTLLIGV** A MFAGSYGKSMFSSKKVINCLPKWRYHFVIA A MFAGSYGKSMFSSKKVINCLPKWRYHFVIA A MFAGSYGKSMFSSKKVINCLPKWRYHFVIA A MFAGSYGKSMFSSKVINCLPKWRYHFVIA A MFAGSYGKSMFSSKVINCLPKWRYHFVIA A MFAGSYGKSMFSSKVINCLPKWRYHFVIA A MFAGSYGKSMFSSKVINCLPKWRYHFVIA A MFAGSYGKSMFSSKVINCLPKWRYHFVIA A MFAGSYGKSMFSSFASAS*VAGTTGVCHHTRLIFVF LV*TGFHYVAQGLELLTA*SYPQLPKVVGL QA VGELHDIFFGNDFDMTFKAQATKEKIDKLN GLMATTYFELLKLSNKLIQ QA VGELHDIFFGNDFDMTFKAQATKEKIDKLN MFAGSYGKSKFTAPKYSSKSSFDDISSSQSN MFAGSYGKSKFTAPKYSSKSSFDDISSSQSN MFAGSYGKSKFTAPKYSSKSSFDDISSSQSN MFAGSYGKSKFTAPKYSSKSSFDDISSSQSN MFAGSYGKSFTAPKYSSKSSFDDISSSQSN MFAGSYGKSFTAPKYSSKSSFDDISSSQSN MFAGSYGKSFTAPKYSSKSSFDDISSSQSN MFAGSYGKSFTAPKYSSKSSFDDISSSQSN MFAGSYGKSFTAPKYSSKSTDTYSS CALCADOROKRINASSASKSLFANNSHVPES RANTATSDVSRKWLIPGAPSSIFTYSDLERA MSKLEAQVKQVEHEMIA SLRHNSRHVRPS RANTATSDVSRKWLIPGAPSSIFTYSDLERA MSKLEAQVKQVEHEMIA SLRHNSRHVRPS RANTATSDVSRKWLIPGAPSSIFTYSDLERA MKSLEAQVKQVEHEMIA SLRHNSRHVRPS RANTATSDVSRKWLIPGAPSSIFTYSDLERA MSKLEAQVKQVEHEMIA SLRHNSRHVRPS LYCKPGMSILSKPLICACHTRITITSSVTAPCPING MKSLEAQVKQVEHEMIA SLRHNSRHVRPS LYCKPGMSILSKPRODQ MFAGGGTTATTTTSSTRAPVSDLT TVPQTQGQTTAQKVSKKTOPRCSTALATGK MFAGSYGFTA MTAGSGRILKRPDQ MFAGGGTTATTTTSSTRAPVSDLT TVP		i	ļ	ļ			
1712 A 3136 1270 274 RVGMVLGTREVCDSTPPSPLYPTFTGREVC HNTWOLSRYVSDLRTDSSNYNPOELWNAG COMV*GGSRDWEGVEEQVEOQVGNKFSSDG VGECSRKLIG*PSVAMSMYDGLWNAG COMV*GGSRDWEGVEEQVEOQVGNKFSSDG VGECSRKLIG*PSVAMSMYDGLKSSTST LEMDICDGHFRQNGGCGYVLKPDFTRIDGSSF HEKRISPPKAQILLQNVISVQLFKVDKTKE GSIVDPLVKVQIFGVPLLTARQETNYVENDS FNPYWGQTLGFRVLGDPPMLRGKMDYDW KSRNDLLGKTPCPGTCMQQGYRHHLLSKDG SIRPASIFYVICQGLEGDES SIRPASIFYVICQGLEGDES APAMSESPLAPHLHQHLVSVPQVLTLIGV** APAMSESPLAPHLHQHLVSVPQVLTLIGV** NFKVFSPTMSPG*PDQSFLQFPPQDLRQFPPDQLRQFPPQDQLRQFPPQQLRQFPPQQRQFQLRQFPPQQRQFQLRQFPPQQRQFQLRQFPPQQRQFQLRQFPPQQRQFQLRQFPPQQRQFQLRQFQPPQQRQFQLRQFPPQQRQFQLRQFPPQQRQFQLRQFPPQQRQFQLRQFQPPQQRQFQLRQFPPQQRQFQLRQFPPQQRQFQLRQFQPPQQRQFQLRQFPPQQRQFQLRQFQPPQQRQFQLRQFQPPQQRQFQLRQFQPPQQRQFQLRQFQPPQQRQFQLRQFQPPQQRQFQLRQFQPQQRQFQLRQFQQRQFQLRQFQQRQFQLRQFQQRQFQLRQFQQRQFQLRQFQQRQFQLRQFQQQRQFQLSQGSKKTQQQLGCQSKKTQQQLGCQSKKTQQLQFQXQPQCRQFQQQQQRQFQLSQGSQSKTQQLGCGSKKTQQQLGCQSKMQQQLGCQSKMQQQLGCQSKMQQQLGCQSKMQQQLGCQSKMQQQLGCQSKMQQQLGCQSCSST SLLKKQQRSTSQTPSQASSQSST SLLKKQQRSTSQTPSQASSQSST SLLKKQQRSTSQTPSQASSQSST SLLKKQQRSTSQTPSQASSQSST SLLKKQQRSTSQTPSQASSQSST SLLKKQQRSTSQTPSQASSQSST SLLKKQQRSTSQTPSQASSQSST SLLKKQQRSTSQTPSQASSQSST SLLKKQQRSTSQTPSQASSQSST SLLKKQQRSTSQTPSQASSQSST SLLKKQQRSTSQTPMARQSST SLKKQQRSTSQTPSQASSQSST SLKKQQRSTSQTPSQASSQSST SLKKQQRSTSQTPSQASSQSST SLKKQQRSTSQTPSQASSQSST SLKKQQRSTSQTPSQASSQSST SLKKQQRSTSQASSQSST SLKKQQRSTSQASSQSST SLKKQQRSTSQASSQSST SLKKQQRSTSQASSQSST SLKKQQRSTSQASSQSST SCKSQSQSQCGTANAQAVVLYPASSQSST SLKKQQRSTSQASSQSST SCKSQSGCGGALVAAGAVVLYPASSQSST SVCQSPASSSQSTGATNAVCSSCXST SVCQSPASSQSLPGASSQSRSRSQSRSRSRSRSQSSQSRST SVCQSPASSSQST SVCQSSTSQASSQSRSRSRSQSSRSRSRSRSQSSQSRST SVCQSSQSST SVCQSSTSQASSQST SVCQSSTSQASSQST SVCQSSTSQASSQST SVCQSSTSQASSQSST SVCQSSTS	1		ł				
HNTWQLSRYYPSDLRTDSSNYNPQELWAG CQMW*GGSRWLGGVEGVEGVEGVEGVEGVEGVEGVEGVEGVEGVEGVEGVE	L	<u>L</u>	<u>i</u> _	L		<u> </u>	
CQMV*GGSRDWEEGVEEQVENKFSSDGT	362	1712	Α	3136	1270	274	RVGMVLGTREVGDSTPPPSPPLYPFTGNEFVQ
VGECSRKLI.G*EMLSVDITSRYRAFSTYLING LEGGIGHGESSSSFILGPSVAMMMQTAGI EMDICDGHFRQNGGCGYVLKPDFLRDIOSSFI EMERISPFRAQTILNQVISVQDLFVDKTKE GSIVDPLVKVQIFGVRLDTARQETNYVENNO FNPYWGGICFRVLGOFFPMLRGKMDYDVENNO FNPYWGGILGRVLGOFFPMLRGKMDYDVENNO FNPYWGGILGRVLGOFFPMLRGKMDYDVENNO FNPYWGGILGRVLGOFFPMLRGKMDYDVENNO FNPYWGGILGRVLGOFFPMLRGKMDYDVENNO FNPYWGGILGRVLGOFFPMLRGKMDYDVENNO SIRPASIFVYLCQEGLEGDES 363							HNTWQLSRVYPSDLRTDSSNYNPQELWNAG
VGECSRKLI.G*EMLSVDITSRYRAFSTYLING LEGGIGHGESSSSFILGPSVAMMMQTAGI EMDICDGHFRQNGGCGYVLKPDFLRDIOSSFI EMERISPFRAQTILNQVISVQDLFVDKTKE GSIVDPLVKVQIFGVRLDTARQETNYVENNO FNPYWGGICFRVLGOFFPMLRGKMDYDVENNO FNPYWGGILGRVLGOFFPMLRGKMDYDVENNO FNPYWGGILGRVLGOFFPMLRGKMDYDVENNO FNPYWGGILGRVLGOFFPMLRGKMDYDVENNO FNPYWGGILGRVLGOFFPMLRGKMDYDVENNO FNPYWGGILGRVLGOFFPMLRGKMDYDVENNO SIRPASIFVYLCQEGLEGDES 363							CQM/V*GGSRDWEEGVEEQQVGNKFSSDGR
LKEGLEGLHGESCSSTLLGPSVAMMMQTAGL	i	ļ		1			VGECSRKLLG*EMLSVDITSRYRAPSTYLLNS
EMDICOGHERQNGGGGYVLKPDELRD(SSF) HPEKPISPEKA, GYLLNQVISVQLPK-VDKTKE GSIVDPLVK VQIFGVELDTARQETNYVENNG FNPYWGQTLCFRV (DIPDFPMLRFGKMDYDW KSRNDLLGKTRCPGTCMQQO FNHIHLLSKND SISLRPASIFVYICIQEGLEGDES SISLRPASIFVYICIQEGLEGDES A							
HPEKPISPEKAQTILLNQVISVQQLPKVDKTRE]	1					
GSIVDPL/VK/QIFG/RLDTARQETHY/VENNG FNPYWGGTL/ERVLQPPPMLRGKMDYDW KSRNDLLGKTPC/GTCMQQGYRHIHLLSKDG ISLRPASIFVYICQEGLEGDES ISLRPASIFVYICQEGLEGDES ISLRPASIFVYICQEGLEGDES ISLRPASIFVYICQEGLEGDES ISLRPASIFVYICQEGLEGDES ISLRPASIFVYICQEGLEGDES ISLRPASIFVYICQEGLEGDES ISLRPASIFVYICQEGLEGDES ISLRPASIFVYICQEGLEGDES ISLRPASIFVYICQEGLEGDES ISLRPASIFVYICQEGLEGDES ISLRPASIFVYICQEGLEGDES ISLRPASIFVYICQEGLEGDES ISLRPASIFVYICQEGLEGT ISLRPASIFVYICQEGLEGT ISLRPASIFVYICQEGLEGT ISLRPASIFVYICQEGLIG ISLRPASIFVYICQEGLIG ISLRPASIFVYICQEGLIG ISLRPASIFVYICQEGLIG ISLRPASIFVO ISLRP	1	1				1	HPEK PISPEK AOTI I NOVISVOOI PK VDKTKE
FNPYWGQTLCFRVLGPDFFMLRFGMDYDW	Į	ļ	j	1			GSIVDDI VK VOIEGVDI DTA DOETNIVVENNIC
SSRNDLIĞKTPCPGTCMQQQYRIIHILISKDG ISLRPASIFYYICIQGELGEQIS ISLRPASIFYYICIQGELGEQIS SILRPASIFYYICIQGELGEQIS SILRPASIFYYICIQGELGEQIS SILRPASIFYYICIQGELGEQIS SILRPASIFYYICIQGELGEQIS SILRPASIFYYICIQGELGEQIS SILRPASIFYYICIQGELGEQIS SILRPASIFY SILLPASIFY SILRPASIFY SILLPASIFY SILRPASIFY SILLPASI				1			
				ļ			
1713 C 3139 60 248	i	l	ł			1	
PAMNESPLAPHLHQHLVFSVFQVLTILIGV**	362	1712	_	1120		240	
1714 A 3140 57 418 SAFKILQLPAFSLYFDLGSLKLILRIHITSIVK NHKVESPRTMSPG*DPQSFLQIPQRPPPQLRV GLTSGLIQHFHSPSSCQFPLLRGPPPFRQPPLGI SGASLCPVLSPPR*PLQPSSL LLPYPSLFVFLRGCHFVTRLECNGVVSAHCN LHLPGSSDSPASAS*VAGTTGVCHHTRLIFVF LV*TGFHYVAQAGLELLTA*SVPQLFKVVGL QA GLEKLHDIRFGNDFDMTPKAQATKEKIDKLN FIKIKKLCIEGYY/NREPQNGRKKT SGASLCPVLSPPR*PQLPKVVGL QA GLEKLHDIRFGNDFDMTPKAQATKEKIDKLN FIKIKKLCIEGYY/NREPQNGRKT SGASLCPVLSPPR*PQLPKVVGL QA GLEKLHORFGNDFDMTPKAQATKEKIDKLN FIKIKKLCIEGYY/NREPQNGRKT SGASLCPVLSPPR*PQLFKVVGL QA GLEKLHORFGNDFDMTPKAQATKEKIDKLN FIKIKKLCIEGYY/NREPQNGRKT SGASLCPVLSPRPPNSKKORIGQSKKT PIEKSDFAAATHPRAFVLSKPDETPNAWMSD SGTGLTYWKLEEKDMHSLPETLEKTFISLSS TDVSPNQVLITLDFILLIMKPRQQISGIQPHGLI NALDDRISS*PDSVLISS*SPSDILDS*SQASN VTSQLPGFPKYPSHTKASPVDSWKNQTFQNE SRTSSTFPSVYTITSSDISVTDEENTTWMVAS ASVSQSQLPGTANSVPECISLTSLEDPVILSKIR QNLKEKHARHIADLRAYYESENSLKQKLEA KEISGVPDWKINTQILVPRCGQLDSALHEATS RVRTLENKNNLLEIEVNDLRERSSAASSASKI LQERIEEMRTSSKEKDNTIIRLKSRLQDLEEAF ENAYKLSDDKEAQLKQENKMFQDLLGEVES LGKEHRRVKDALNITENKLLDAYTQISDLKR MISKLEAQVQVQVEHRINSKLENSSPISTSST SLLIKKQRESTSDTPMRALKELDEGKIFNWG TQTEKEDTSNSLAYNWEDENSTVASSSPEK CAQQRKRLNSASQRSSLPPSNRKSSTPTKR EIMLTPVTVAYSPKRSPKENLSPGFSHLLSKN ESSPIREKTYSEKATDNHVNISSCPEPVPNGV KKVSVRTAWERNKSVSPGCKPVSVTPGN DFEYTAKIRTLAETERFFDELTKEKDQIEAAL SRMPSPGGRITLQTRNQVKCLSINLL SRMPSPGGRITLQTRNQVKCLSINLL SRMPSPGGRITLQTRNQVKCLSINLL SRMPSPGGRITLQTRNQVKCLSINLL SRMPSPGGRITLQTRNQVKCLSINLL SRMPSPGGRITLQTRNQVKCLSINLL NAGRET RVPGSPAPSSLPLHSPGACGTEVDMDPQRSPL LEVKGNIELKRPLIKAPSQLPJ.SGSRLKRPDQ MEDGLEPEKKRTIKASQLPJ.SGSRLKRPDQ MEDGLEPEKKRTRGLGATTKITTSHPRVPSLT TVPQTGQQTTAQKVSKKTGPRCSTAIATGLK NQKVPAPVQKSCTSGVPMAGKKKFSKPR NQKVPAPVQKSCTSGVPMAGKKFSKPR NQKVPAPVQKSCTSGVPMAGKKFSKPR NQKVPAPVQKSCTSGVPMAGKKFSKPR NQKVPAPVQKSCTSGVPMAGKKFSKPR NQKVPAPVQKSCTSGVPMAGKKFSKPR NQKVPAPVQKSCTSGVPMAGKKFSKPR NQKVPAPVQKSCTSGTATATGLK NQKVPAPVQKSCTSGTATATGLK NQKVPAPVQKSCTSGTATATGLK NQKVPAPVQKSCTSGTATATGLK NQKVPAPVQKSCT	363	1/13	C	3139	60	248	
NHKVESPRTMSPG*DPQSFLQIPQPRPPQLRV GLTSGLIQHFHSPSSCQFPLLRGPPPRQPPLGI SGASLCPULSPPR*PLQPSSL 1715 A 3145 122 413 LLPYPSLFVFLRQCHFVTRLECKGVVSAHCN LHPGSSDSPASAS*VAGTTGVCHHTRLIFLVF LV*TGFHYVAQAGLELLTA*SVPQLPKVVGL QA 366 1716 A 3150 247 2 VGEKLHDIRFGNDFDMTPKAQATKEKIDKLN FIKIKKLCIEGYYNREPQNGRKIFANYVSDK GLMATIYEELLKLSNKLIQ 367 1717 A 3152 3 2367 QKLKQNQPKRAHVEDGGSRSKQGNEQSKKT PIEKSDFAAATHPRAFYLSKPDETPNAWMSD SGTGLTVWKLEEKDMISJEPTLEKTFISLSS TDVSPNQVLTLDPTLHMKPKQQISGIQPHGLP NALDDRISFSPDSVLEPSMSPSDIDSFQASN VTSQLPGFPKYPSHTKASSYDDSWKNQTFQNE SRTSSTFPSVYTITSNDISVNTVDEENTVMVAS ASVSQSQLPGTANSVPSCISLTSLEDPVLISKIR QNLKEKHARHLADLRAYYESEINSLKQKLEA KEISGVEDWKITNQILVDRCGQLDSALHEATS RVKTLENKNILLEIEVNDLRERFSAASSASKI LQERIEEMRTSSKEKDNTITRLKSRLQDLEGAF ENAYKLSDDKEAQLKORKMFQDLLGGYES LGKEHRRVKDALNTTENKLLDAYTQISDLKR MISKLEAQVKQVEHENNLSLERHNSRIHVRFS RANTLATSDVSRKKWLIPGAEYSIFTGQPLDT QDSNVDNQLEETCSLGHRSFLEKDGSPGSSST SLLIKKQRETSDTPIMRLKELDEGKIFNWG TQTEKEDTSNSLL*/INPRQTETSVNASRSPEK CAQQRQKRLINSASQRSSSLPFSNRKSSTPTKR EIMLTFVTVAYSPKRSPLSFSFILLSKN ESSPIREKTYSEKATDNHVNHSSCPEFVPNGV KKYSKTAWEKNKSVSPEQCKPVSVTPQN DFEYTAKIRTLAETERFFDELTKEKDQIEAAL SRMPSPGGRTILQTRLNQVKCLSINLL 368 1718 A 3163 2 2350 EFKSGGGGGGICVAAGAVLVLYYPASRAGERT RVPGSPAPSSLPLHSPGAGGLEVSVTPQN MEDGLEPEKKRTRGLGATTKITTSHPRVPSLT TVPQTQGQTTAQKVSKKTGPRCSTAIATGLK NQKFVPAVPQVSKSTGSVPMAGGKKFSKPR							
GLTSGLIQHFHSPSSCQFPLLRGPFPRQPPLGI SGASLCPVLSPR-PLQPSSL 365 1715 A 3145 122 413 LLPYPSLFVFLRQCHFVTRLECNGVVSAHCN LH.PGSSDSPASAS*VAGTTGVCHHTRLIFVF LV*TGFHYVAQAGLELLTA*SVPPQLPKVVGL QA 366 1716 A 3150 247 2 VGEKLHDIRFGNDFDMTPKAQATKEKIDKLN FIKIKKLCIEGYYJNREPQNGRKIFANYVSDK GLMATTYEELKLSNKLIQ 367 1717 A 3152 3 2367 QKLKQNQPKRAHVEDGGSRSKQGNEQSKKT PIEKSDFAAATHPRAFYLSKPDETPNAWMSD SGTGLTYWKLEEKDMHHSLPTLEKTFISLSS TDVSPNQV.TLDPTLHMKPKQQISGIQPHGLP NALDDRISFSPDSVLEPSMSSPSDIDSFSQASN VTSQLPGFPKYPSHTKASPYDSWKNQTFONE SRTSSTFSVYTITSNDISVNTVDEENTVMVAS ASVSQSQLPGTANSVPECISLTSLEDPVILSKIR QNLKEKHARHIADLRAYYESEINSLKQKLEA KEISGVEDWKITNQILVDRCGQLDSALHEATS RVRTLENKNNLLEIEVNDLRERFSAASSASKI LQERIEBMTSSKEKDNTIRLKSRLQDLEEAF ENAYKLSDDKEAQLKQENKMFQDLLGEYES LGKEHRRVKDALNTTENKLLDAYTQISDLKR MISKLEAQVKQVEHEMMLSLRHNSRIHVRPS RANTLATSDVSRRKWLIPGAEVSIFTGQPLDT QDSNVDNQLEETCSLGHRSPLEKDSSP/GSSST SLLIKKQRETSDTPIMRAKELDEGKIFKNWG TQTEKEDTSNSLL*/INPRQTIETSVNASRSPEK CAQQRQKRINNSASQRSSSLPPSNRKSSTPTKR EIMLTPVTVAYSPKRSPKENLSPGFSHLLSKN ESSPIREKTYSKATDHNVHNISCPEPVPNGV KKVSVRTAWEKNSVSYEQCKPVSVTPQGN DFEYTAKIRTLAETERFFDELTKEKDQIEAAL SRMPSPGGRTLQTRLNQVKCLSLNLL 368 1718 A 3163 2 2350 EFKSGGGGGGLVAAGAVLVLYPASRAGERT RVPGSPAPSSLPLHSRGACGTEVDMDPQRSFL LEVKGNIELKRPLIKAPSQLPLSGSRIKRRPDQ MEDGLEPEKKTTRGLGATTKITTSHPRVPSLT TVPQTQGQTTAQKVSKKTOPRCSTALATGLK NQKPVPAVPVVKSCSTVPPMAGRKFSKRFPRD	364	1714	ļ A.	3140	57	418	
GLTSGLIQHFHSPSSCQFPLLRGPFPRQPPLGI SGASLCPVLSPR-PLQPSSL 365 1715 A 3145 122 413 LLPYPSLFVFLRQCHFVTRLECNGVVSAHCN LH.PGSSDSPASAS*VAGTTGVCHHTRLIFVF LV*TGFHYVAQAGLELLTA*SVPPQLPKVVGL QA 366 1716 A 3150 247 2 VGEKLHDIRFGNDFDMTPKAQATKEKIDKLN FIKIKKLCIEGYYJNREPQNGRKIFANYVSDK GLMATTYEELKLSNKLIQ 367 1717 A 3152 3 2367 QKLKQNQPKRAHVEDGGSRSKQGNEQSKKT PIEKSDFAAATHPRAFYLSKPDETPNAWMSD SGTGLTYWKLEEKDMHHSLPTLEKTFISLSS TDVSPNQV.TLDPTLHMKPKQQISGIQPHGLP NALDDRISFSPDSVLEPSMSSPSDIDSFSQASN VTSQLPGFPKYPSHTKASPYDSWKNQTFONE SRTSSTFSVYTITSNDISVNTVDEENTVMVAS ASVSQSQLPGTANSVPECISLTSLEDPVILSKIR QNLKEKHARHIADLRAYYESEINSLKQKLEA KEISGVEDWKITNQILVDRCGQLDSALHEATS RVRTLENKNNLLEIEVNDLRERFSAASSASKI LQERIEBMTSSKEKDNTIRLKSRLQDLEEAF ENAYKLSDDKEAQLKQENKMFQDLLGEYES LGKEHRRVKDALNTTENKLLDAYTQISDLKR MISKLEAQVKQVEHEMMLSLRHNSRIHVRPS RANTLATSDVSRRKWLIPGAEVSIFTGQPLDT QDSNVDNQLEETCSLGHRSPLEKDSSP/GSSST SLLIKKQRETSDTPIMRAKELDEGKIFKNWG TQTEKEDTSNSLL*/INPRQTIETSVNASRSPEK CAQQRQKRINNSASQRSSSLPPSNRKSSTPTKR EIMLTPVTVAYSPKRSPKENLSPGFSHLLSKN ESSPIREKTYSKATDHNVHNISCPEPVPNGV KKVSVRTAWEKNSVSYEQCKPVSVTPQGN DFEYTAKIRTLAETERFFDELTKEKDQIEAAL SRMPSPGGRTLQTRLNQVKCLSLNLL 368 1718 A 3163 2 2350 EFKSGGGGGGLVAAGAVLVLYPASRAGERT RVPGSPAPSSLPLHSRGACGTEVDMDPQRSFL LEVKGNIELKRPLIKAPSQLPLSGSRIKRRPDQ MEDGLEPEKKTTRGLGATTKITTSHPRVPSLT TVPQTQGQTTAQKVSKKTOPRCSTALATGLK NQKPVPAVPVVKSCSTVPPMAGRKFSKRFPRD		1	l				NHKVESPRTMSPG*DPQSFLQIPQPRPPQLRV
SGASLCPULSPRPPLQPSSL		1	İ				GLTSGLIQHFHSPSSCQFPLLRGPPFPRQPPLGI
365 1715 A 3145 122 413 LLPYPSLFVELRQCHFVTNLECNGVVSAHCN LHLPGSSDSPASAS+VAGTTGVCHHTRLIF\VF LV*TGFHYVAQAGLELLTA*SYPQLPKVVGL QA 366 1716 A 3150 247 2 VGEKLHDIRFGNDFDMTPKAQATKEKIDKLN FIKIKKLCIEGYY/NREPQNGRKIFANTVSUDK GLMATIYEELLKLSNKLIQ GLMATIYEELLKLSNKLIQ GKLKONQPKRAHVEGGGRSKQGNEQSKKT PIEKSDFAAATHPRAFYLSKPDETPNAWMSD SGTGLTYWKLEEKDMHHSLETLEKTFISLSS TDVSFNQVLTLDPTLHMKPKQQISGIPHGLP NALDDRISFSPDSVLEPSMSSPDIDSFSQASN VTSQLPGFPKYPSHTKASPYDSWKNQTFQNE SRTSSTFPSVYTITSDISVNTVDEENTVMVAS ASVSQSQLPGTANSVPECISLTSLEDPVILSKIR QNLKEKHARHIADLRAYYESIENSLKQKLEA KEISGVEDWKITNQILVDRCGQLDSALHEATS RVRTLENKNNLLEIEVNDLERFSAASSASKI LQERIEEMTSSKEKSDNTIRLKSRLQDLEAF ENAYKLSDDKEAQLKQENKMPQDLLGEYES LGKEHRRVKDALNTTENKLLDAYTQISDLKR MISKLEAQVKQVEHEMMLSLRHNSRIHVRPS RANTLATSDVSRKWLIPGAEVSIFTGQPLDT QDSNVDNQLEETCSHRSPLEKDSSPGSSST SLIKKQRETSDTPIMRALKELDEGKIFKNWG TQTEKEDTSNSLL*/INPRQIETSVNASRSPEK CAQQRQKRINNSASQRSSSLPPSNRKSSTPTKR EIMLTPVTVAYSPKRSPKENLSPGFSHLLSKN ESSPIREKTYSEKATDNHVNHSSCPEPVPRIGV KKVSVRTAWEKNKSVSYEQCKPVSVTPQGN DFEYTAKIRTLAETERFFDELTKEKDQIEAAL SRMPSPGGRTILQTRLNQVKCLSLNLL 368 1718 A 3163 2 2350 EFKSGGGGAGLVAGAVLVLYPASRAGERT RVPGSPAPSSLPLHSPGACGTEVDMDPQRSFL LEVKGNIELKRPLIKSGPLPLSGGRIKRRPDQ MEDGLEPEKRTRGLGATTKITTSHRVPSLT TVPQTQCQTTAQKVSSKKTOPRCSTALATGLK NQKPVPAYPVPKVSKGTSVPPMAGRKFSKRP	l	ł	}	1 .		1	
LHLPGSSD\$PASA\$*VAGTTGVCHHTRLIFWF LV*TGFHYVAQAGLELTA*SWPQLPKVVGL QA 366 1716 A 3150 247 2 VGEKLHDIRFGNDFDMTPKAQATKEKIDKLN FIKIKKLCIEGYY/NREPQNGRKIFANYVSUDK GLMATTYEELKISNKLIQ 367 1717 A 3152 3 2367 OKLKQNOPKRAHVEDGGSR\$KQGNEQ\$KKT PIEKSDFAAATHPRAFYLSKPDETPNAWMSD SGTGLTYWKLEEKDMHHSI.PETLEKTFISLS\$ TDV\$PNQVLTLDPTLHMKPKQQISGIQPHGLP NAL DDRISS*PSDVL*PSMSSP\$DIDSFSQASN VTSQLPGFPK YPSHTKASPVDSWKNQTFQNE SRTSSTFPSVYTITISNDISVNTVDEENTVMVAS ASVSQQLPGTANSVPECISLTSLEDPVII.SKIR QNLKEKHARHIADLRAYYESEINSLKQKLEA KEISGVEDWKITNQIL VDRCQGLDSALHEATS RVRTLENKNNILLEIEVNDLRERFSAASSASKI LQERIEEMRTSSKEKDNTIRLKSRLQDLEEAF ENAYKLSDDKEAQLKQENKMFQDLLGEYES LGKEHRRVKDALNTTENKLLDAYTQISDLKR MISKLEAQVKQVEHENMI.SLRENSRHVRPS RANTLATSDVSRRKWLIPGAEYSIFTGQPLDT QDSNVDNQLEETCSLGHRSPLEKDSSP/GSSST SLLIKKQRETSDTPIMRALKELDEGKIFKNWG TQTEKDTSNSLL*INPRQTETSVNASRSPEK CAQQRKRLNSASQRSSSLPPSNRKSSTPTKR EIMLTPVTVAYSPKRSPKENLSPGFSHLLSKN ESSPIREKTYSEKATDNHVNHSSCPPVPNGV KKVSVRTAWEKNKSVSYEQCKPVSVTPQGN DFEYTAKIRTLAETERFFDELTKEKDQIEAAL SRMPSPGGRITJORDAT SRMPSGRITJORDAT RVPGSPAPSSLPLHSPGAGGTEVDMDPQRSPL LEVKGORIELKREKLKAPSQLPLSGSRLKRRPDQ MEDGLEPEKKRTRGLGATTKITTSHRVPSLT TVPQTTQGQTTAAQKVSKTGPRCSTAIATGLK NQKPVPAVPVQKSGTSGVPPMAGGGKFSKRP	365	1715	Α	3145	122	413	
366 1716 A 3150 247 2 VGEKLHDIRFGNDFDMTPKAQATKEKIDKLN FIKIKKLCEGYYMREPQNGRKIFANYVSUDK GLMATTYELLKLSNKLD, 367 1717 A 3152 3 2367 OKLKQNOPKRAHVEDGGSRSKQGNEQSKKT PIEKSDFAAATHPRAFYLSKPDETPNAWMSD SGTGLTYWKLEEKDMIHSLPETLEKTISLSS TDVSPRQVLTLDPTLHMKPKQQISGIQPHGLP NALDDRISFSPDSVLEPSMSPSDIDSFSQASN VTSQLPGFPKYPSHTKASPVDSWKNQTFQNE SRTSSTFPSVYITTSNDISVNTVDEENTVMVAS ASVSQSQLPGTANSVPECISLTSLEDPVILSKIR QNLKEKHARHIADLRAYYESEINSLKQKLEA KEISGVEDWKITNQILVDRCGGQLDSALHEATS RVRTLENKNNLLEIEVNDLRERFSAASSASKI LQERIEEMRTSSKEKDNTIIRLKSRLQDLEAF ENAYKLSDDKEAQLKQENKMFQDLLGEVES LGKEHRRVKDALNTTENKLDAYTQISDLKR MISKLEAQVKQVEHEMLSLRHNSRIHVRPS RANTLATSDVSRRKWLIPGAEYSEFTGQPLDT QDSNVDNQLEETCSLGHRSPLEKDSSPGSSST SLLIKKQRETSDTPIMRALKELDEGKIFKNWG TQTEKEDTSNSLL*INPRQTETSVNASRSPEK CAQQRQKRLNSASORSSLPSNRKSSTPTKR EIMLTVYVYSPKRSPKENLSPGFSHLLSKN ESSPIREKTYSEKATDNHVNHSSCPEPVPNGV KKVSVRTAWERNKSVSYEQCKPVSVTPQGN DFEYTAKIRTLAETERFFDELTKEKDQIEAAL SRMPSPGGRITLQTILNQVKCLSLNLL SRMPSPGGRITLQTILNQVKCLSLNLL SRMPSPGGRITLQTILNQVKCLSLNLL SRMPSPGGRITLQTILNQVKCLSLNLL SRMPSPGGRITLQTILNQVKCLSLNLL VRYPGSPAPSSLPLHSPGAGGTEVDMDPQRSPL LEVKGNIELKRFRIKAPSQLPLSGSRLKRRPDQ MEDGLEPEKKRTRGLGATTKITTSHPRYPSLT TVPQTQGQTTAQKVSKTGPRCSTAIATGLK NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP		i					
366 1716 A 3150 247 2 VGEKLHDIRFGNDFDMTPKAQATKEKIDKLN FIKIKKLCIEGYY/NREPONGRKIFANYVSUDK GLMATIYEELLKLSNKLIQ 367 1717 A 3152 3 2367 QKLKQNQPKRAHVEDGGRSKQGNEQSKKT PIEKSDFAAATHPRAFYLSKPDETPNAWMSD SGTGLTYWKLEEKDMHHSLPETLEKTFISLSS TDVSPNQVLTLDPTLHMKPKQQISQIQPHGLP NALDDRISFSPDSVLEPSMSSPSDIDISFSQASN VTSQLPGFPKYPSHTKASPVDSWKNQTFQNE SRTSSTFPSVJTITSNDISVNTVDEENTVMVAS ASVSQSQLPGTANSVPECISLTSLEDPVILSKIR QNI.KEKHARHIADLRAYYESEINSLKQKLEA KEISGYEDWKITNQILVDRCGQLDSALHEATS RVRTLENKNNLLEIEVNDLRERFSAASSASKI LQERIEEMRTSSKEKDTITIRLKSRLQDLEEAF ENAKLSDDKEAQLKQENKMFQDLLGEVSE LGKEHRRVKDALNTTENKLLDAYTQISDLKR MISKLEAQVKQVEHENMI-SILRHNSRIHVRPS RANTLATSDVSRKWLIPGAEYSIFTGQPLDT QDSNVDNQLEETCSLGHRSPLEKDSSPGSSST SILLIKKQRETSDTPIMRALKELDEGKIFKNWG TQTEKEDTSNSL*NIPRQTETSVNASRSPEK CAQQRQKRLNSASQRSSSLPPSNRKSSTPTKR EIMLTPVTVAYSPKRSKENLSPGFSHLLSKN ESSPIREKTYSEKATDNHVNHSCPEPVPNGV KKVSVRTAWEKNKSVSYEQCKIPVSVTPQGN DFEYTAKIRTLAETERFDELTKEKQUIEAAL SRMPSPGGRITLQTRLNOVKCISINLL SRMPSPGGRITLQTRLNOVKCISINLL SRMPSPGGRITLQTRLNOVKCISINLL LEVKGNIELKRIPLKAPSQLPLSGSRILKRRPDQ MEDGLFFEKKRTIGLGATTKITTSHPVPSLT TVPQTQGQTTAQKVSKKTGPRCSTAIATGLK NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP		ŀ	l				
366 1716 A 3150 247 2 VGEKLHDIRFGNDFDMTPKAQATKEKIDKLN FIKIKKLCIEGYYJNREPONGRKIFANYVSUK GLMATIYEELIKLSNKLIQ 367 1717 A 3152 3 2367 QKLKQNQPKRAHVEDGGSRSKQGNEQSKKT PIEKSDFAAATHPRAFYLSKPDETPNAWMSD SGTGLTYWKLEEKDMHSLPETLEKTFISLSS TDVSPNQVLTLDPTLHMKPKQQISGIQPHGLP NAALDRISFSPDSVLEPSMSSPSDIDSFSQASN VTSQLPGFPKYPSHTKASPVDSWKNQTFONE SRTSSTFPSVYTITSNDISVNTVDEENTVMVAS ASVSQSQLPGTANSVYPECISLTSLEDPVILSKIR QNLKEKHARHADLRAYYESEINSLKQKLEA KEISGVEDWKITNQILVDRCGQLDSALHEATS RVRTLENKNNLLEIEVNDLRERFSAASSASKI LOERIEFMRTSKEKDNTTIRLKSRLQDLEEAF ENAYKLSDDKEAQLKQENKMFQDLLGEYES LGKEHRRVKDALNTTENKLLDAYTQISDLKR MISKLEAQVKQHENMIS.IRHNSRIHVRPS RANTLATSDVSRRKWLIPGAEYSIFTGQPLDT QDSNVDNQLEETCS.IGHRSPLEKDSSPGSSTS SLLIKKQRETSDTPIMRALKELDEGKIFKNWG TQTEKEDTSNSLL*/INPRQTETSVNAASRSPEK CAQQRQKRLNSASQRSSSLPSNRKSSTTTKR EIMLTPVTVAYSPKRSPKENLSPGFSHLLSKN ESSPIREKTYSEKATDNHVNHSSCPEPVPNGV KKVSVRTAWEKNKSVSYEQCKPVSVTPQGN DFEYTAKIRTLAETERFFDELTKEKDQIEAAL SRMPSPGGRITLQTRLNQVKCLSLNLL SRMPSPGGGAGLVAAGAVLVLYPASRAGERT RVPGSPAPSSLPLHSPGACGTEVVDTPQRSPL LEVKGNIELKRPLIKAPSQLPLSGSRLKRRPDQ MEDGLEFEKKRTIGLGATTKITTSHPRVPSLT TVPQTQCQTTAQKVSKKTGPRCSTAIATGLK NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP	1	ļ		j [
FIKIKKLCIEGYY/NREPQNGRKIFANYVS/DK GLMATIYEELLKLSNKLIQ 367 1717 A 3152 3 2367 QKLKQNQPKRAHVEDGGRSKQGNEQSKKT PIEKSDFAAATHRAFYLSKPDETPNAWMSD SGTGLTYWKLEEKDMHHSLPETLEKTFISLSS TDVSPNQVLTLDPTLHMKPKQQISGIQPHGLP NALDDRISFSPDSVLEPSMSSPSDIDISFSQASN VTSQLPGFPKYPSHTKASPVDSWKNQTFQNE SRTSSTFPSVYTITSNDISVNTVDEENTVMVAS ASVSQSQLPGTANSVPECISLTSLEDPVILSKIR QNLKEKHARHIADLRAYYESEINSLKQKLEA KEISGVEDWKITNQIL VDRCGQLDSALHEATS RVRTLENKNNLLEIEVNDLRERFSAASSASKI LQERIEEMRTSSKEKDNTIIRLKSRLQDLEEAF ENAYKLSDDKEAQLKQENKMFQDLLGEYES LGKEHRRVKDALNTTENKLLDAYTQISDLKR MISKLEAQVKQVEHEMMLSLRHNSRIHVRPS RANTLATSDVSRRKWLIPGAEYSIFTGQPLDT QDSNVDNQLEETCSLGHRSPLEKDSSP/GSSST SLLIKKQRETSDTPIMRALKELDEGKIFKNWG TQTEKEDTSNSLL*/INPRQTETSVNASRSPEK CAQQRQKRLNSASQRSSSLPPSNRKSSTPTKR EIMLTPVTVAYSPKRSPKENLSPGFSHLLSKN ESSPIREKTYSEKATDNHVNHSSCPEPVPNGV KKYSVRTAWEKNKSSVYEQKPVSVTPQGN DFEYTAKIRTLAETERFFDELTKEKDQIEAAL SRMPSPGGRITLQTRLNQVKCLSLNLL SRMPSPGGRITLQTRLNQVKCLSLNLL LEVKGNIELKRPLIKAPSQLPLSGSRLKRRPDQ MEDGLFPEKKRTIGLGATTKITTSHRVPSLT TVPQTQGQTTAQKVSKKTGPRCSTAIATGLK NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP	366	1716	Δ	3150	247	2	
GLMATTYEELLKLSNKLIQ 367 1717 A 3152 3 2367 QKLKQNQPKRAHVEDGGSRSKQGNEQSKKT PIEKSDFAAATHPRAFYLSKPDETPNAWMSD SGTGLTYWKLEEKDMHHSLPETLEKTFISLSS TDVSPNQVLTLDPTLHMKPKQQISGIQPHGLP NALDDRISFSPDSVLEPSMSSPDIDDFSQASN VTSQLPGFPKYPSHTKASPVDSWKNQTFQNE SRTSSTFPSVYTITISNDISVNTVDEENTVMVAS ASVSQSQLPGTANSVPECISLTSLEDPVILSKIR QNLKEKHARHIADLRAYYESEINSLKQKLEA KEISGVEDWKITNQILVDRCGQLDSALHEATS RVRTLENKNNLLEIEVNDLREFFSAASSASKI LQERIEBRTSSKEKDNTIIRLKSRLQDLEEAF ENAYKLSDDKEAQLKQENKMFQDLLGEYES LGKEHRRVKDALNTTENKLLDAYTQISDLKR MISKLEAQVKQVEHENMLSLRHNSRIHVRPS RANTLATSDVSRRKWLIPGAEVSIFTGQPLDT QDSNVDNQLEETCSLGHRSPLEKDSSPJOSSST SLLIKKQRETSDTPIMRALKELDEGKIFKNWG TQTEKEDTSNSL1*/INPRQTETSVNASRSPEK CAQQRQKRLNSASQRSSSLPPSNRKSSTPTKR EIMLTPVTVAYSPKRSPKENLSPGFSHLLSKN ESSPIREKTYSEKATDNHVNHSSCPEPVPNGV KKVSVRTAWERNKSVSYEQCKPVSVTPQGN DFEYTAKIT.AETERFFDELTKEKDQIEAAL SRMPSPGGRITLQTRLNQVKCLSLNLL 368 1718 A 3163 2 2350 EFKSGGCGGLVAAGAVLVLYPASRAGERT RVPGSPAPSSLPLHSPACGTEVDMDPQRSPL LEVKGNIELKRPLKAPSQLPLSGSRLKRRPDQ MEDGLEPEKKRTRGLGATTKITTSHPRVPSLT TVPQTQGQTTAQKVSKKTGPRCSTAIATGLK NQKPVPAPVQKSGTSGVPPMAGGKKPSKRP	300	1710	Α	3130	247		
367 1717 A 3152 3 2367 QKLKQNQPKRAHVEDGGSRSKQGNEQSKKT PIEKSDFAATHPRAYLSKPDETPNAWMSD SGTGLTYWKLEEKDMHRSLPETLEKTISLSS TDVSPNQVLTLDPTLHMKPKQQISGIQPHGLP NALDDRISFSPDSVLEPSMSSPSDIDSFSQASN VTSQLPGFPKYPSHTKASPVDSWKNQTFQNE SRTSSTFPSVTITISNDISVNTVDEENTVMVAS ASVSQSQLPGTANSVPECISLTSLEDPVILSKIR QNLKEKHARHIADLRAYYESEINSLKQKLEA KEISGVEDWKITNQILVDRCGQLDSALHEATS RVRTLENKNNLLEIEVNDLRERFSAASSASKI LQERIEEMRTSSKEKDNTIIRLKSRLQDLEEAF ENAYKLSDDKEAQLKQENKMFQDLLGEYES LGKEHRRYKDALNTTENKLLDAYTQISDLKR MISKLEAQVKQVEHENMLSLRHNSRIHVRPS RANTLATSDVSRRKWLIPGAEYSIFTGQPLDT QDSNVDNQLEETCSLGHRSPLEKDSSP/GSSST SLLIKKQRETSDTPIMRALKELDEGKIFKNWG TQTEKEDTSNSLL*INPRQTETSVNASRSPEK CAQQRQKRLNSASQRSSLPPSNRKSSTPTKR EIMLTPVTVAYSPKRSPKENLSPGFSHLLSKN ESSPIREKTYSEKATDNHVNHISSCPEPVPNGV KKVSVRTAWEKNKSVSYEQKPVSVTPQGN DFEYTAKIRTLAETERFFDELTKEKDQIEAAL SRMPSPGGRITLQTRLNQVKCLSLNLL SRMPSPGGRITLQTRLNQVKCLSLNLL SRMPSPGGRITLQTRLNQVKCLSLNLL LEVKGNIELKRPLGGTEVDMDPQRSPL LEVKGNIELKRPLGGTEVDMDPQRSPL LEVKGNIELKRPLGGTEVDMDPQRSPL LEVKGNIELKRPLGGTEVDMDPQRSPL LEVKGNIELKRPLGGTETATGLK NQKPVPAVPQKSGTISGVPPMAGGKKPSKRP	l		ł	1			
PIEKSDFAAATHPRAFYLSKPDETPNAWMSD SGTGLTYWKLEERDMHHSLPETLEKTFISLSS TDVSPNQVLTLDPTLHMKPKQQISGIQPHGLP NALDDRISFSPDSVLEPSMSSPSDIDSFSQASN VTSQLPGFPKYPSHTKASPVDSWKNQTFQNE SRTSSTFFSVYTITSNDISVNTVDEENTVMVASS ASVSQQLPGTANSVPECISLTSLEDPVILSKIR QNLKEEKHARHIADLRAYYESEINSLKQKLEA KEISGVEDWKITNQIL VDRCGQLDSALHEATS RVRTLENKNNLLEIEVNDLEREFSSASSASKI LQERIEEMRTSSKEKDNTIIRLKSRLQDLEEAF ENAYKLSDDKEAQLKQENKMFQDLLGEYES LGKEHRRVKDALNTTENKLLDAYTQISDLKR MISKLEAQVKQVEHENMLSLRHNSRIHVRPS RANTLATSDVSRRKWLIPGAEYSIFTGQPLDT QDSNVDNQLEETCSLGHRSPLEKDSSP/GSSST SILIKKQRETSDTPIMRALKELDEGKIFKNWG TQTEKEDTSNSLL*/INPRQTETSVNASRSPEK CAQQRQKRLNSASQRSSSLPPSNRKSSTPTKR EIMLTPVTVAYSPKRSPKENLSPGFSHLLSKN ESSPIREKTYSEKATDNHVNHSSCPEPVPNGV KKVSVRTAWEKNKSVYEQCKPVSVTPQGN DFEYTAKIRTLAETERFFDELTKEKDQIEAAL SRMPSPGGRITLQTRLNQVKCLSLNLL 368 1718 A 3163 2 2350 EFKSGGCGAGLVAAGAVLVLYPASRAGERT RVPGSPAPSSLPLHSPGACGTEVDMDPQRSPL LEVKGNIELKRPLIKAPSQLPLSGSRLKRRPDQ MEDGLEPEKKRTRGLGATTKKITTSHPRVPSLT TVPQTQGQTTAQKVVSKKTGPRCSTAIATGLK NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP	367	1717	<u> </u>	21.50		00.6	GLMATIYEELLKLSNKLIQ
SGTGLTYWKLEEKDMHHSLPETLEKTFISLSS TDVSPNQVLTLDPTLHMKPKQQISGIQPHGLP NAL DDRISFSPDSVLEPSMSSPSDIDSFSQASN VTSQLPGFPKYPSHTKASPVDSWKNQTFQNE SRTSSTFPSVYTITSNDISVNTVDEENTVMVAS ASVSQSQLPGTANSVPECISLTSLEDPVILSKIR QNLKEEKHARHIADLRAYYESEINSLKQKLEA KEISGVEDWKITNQILVDRCGQLDSALHEATS RVRTLENKNNLLEIEVNDLRERFSSASSASKI LQERIEEMRTSSKEKDNTIIRLKSRLQDLEEAF ENAYKLSDDKEAQLKQENKMFQDLLGEYES LGKEHRRVKDALNTTENKLLDAYTQISDLKR MISKLEAQ VKQVEHENMLSLRHNSRIHVRPS RANTLATSDVSRRKWLIPGAEYSIFTGQPLDT QDSNVDNQLEETCSLGHRSPLEKDSSP/GSSST SLLIKKQRETSDTPIMRALKELDEGKIFKNWG TQTEKEDTSNSLL*/INPRQTETSVNASRSPEK CAQQRQKRLNSASQRSSSLPPSNRKSSTPTKR EIMLTPVTVAYSPKRSPKENLSPGFSHLLSKN ESSPIREKTYSEKATDNHVNHSSCPEPVPNGV KKVSVRTAWEKNKSVSYEQCKPVSVTPQGN DFEYTAKIRTLAETERFFDELTKEKDQIEAAL SRMPSPGGRITLQTRLNQVKCLSLNLL 368 1718 A 3163 2 2350 EFKSGGCGAGLVAAGAVLVLYPASRAGERT RVPGSPAPSSLPLHSPGACGTEVDMDPQRSPL LEVKGNIELKRPLIKAPSQLPLSGSRLKRRPDQ MEDGLEPEKKRTRGLGATTKKITTSHPRVPSLT TVPQTQGQTTAQKVSKKTGPRCSTAIATGLK NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP	307	1/1/	A	3132	3	2367	
TDVSPNQVLTLDPTLHMKPKQQISGIQPHGLP NALDDRISFSPDSVLEPSMSSPSDIDSFSQASN VTSQLPGFPKYPSHTKASPVDSWKNQTFQNE SRTSSTFPSVYTITSNDISVNTVDEENTVMVAS ASVSQSQLPGTANSVPECISLTSLEDPVILSKIR QNLKEKHARHIADLRAYYESEINSLKQKLEA KEISGVEDWKITNQIL VDRCGQLDSALHEATS RVRTLENKNNLLEIEVNDLRERFSAASSASKI LQERIEEMRTSSKEKDNTIIRLKSRLQDLEEAF ENAYKLSDDKEAQLKQEDKMFQDLLGEYES LGKEHRRVKDALNITENKLLDAYTQISDLKR MISKLEAQVKQVEHENMLSLRHNSRIHVRPS RANTLATSDVSRRKWLIPGAEYSIFTGQPLDT QDSNVDNQLEETCSLGHRSPLEKDSSP/GSSST SLLIKKQRETSDTPIMRALKELDEGKIFKNWG TQTEKEDTSNSLL*/INPRQTETSVNASRSPEK CAQQRQKRLNSASQRSSSLPPSNRKSSTPTKR EIMLTPVTVAYSPKRSPKENLSPGFSHLLSKN ESSPIREKTYSEKATDNHVNHSSCPEPVPNGV KKVSVRTAWEKNKSVSYEQCKPVSVTPQGN DFEYTAKIRTLAETERFFDELTKEKDQIEAAL SRMPSPGGRITLQTRLNQVKCLSLNLL 368 1718 A 3163 2 2350 EFKSGGCGAGLVAAGAVLVLYPASRAGERT RVPGSPAPSSLPLHSPGACGTEVDMDPQRSPL LEVKGNIELKRPLIKAPSQLPLSGSRLKRRPDQ MEDGLEPEKKRTRGLGATTKITTSHPRVPSIT TVPQTQGQTTAQKVSKKTGPRCSTAIATGLK NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP		·					
NALDDRISFSPDSVLEPSMSSPSDIDSFSQASN VTSQLPGFPKYPSHTKASPVDSWKNQTFQNE SRTSSTFPSVYTITSNDISVNTVDEENTVMVAS ASVSQSQLPGTANSVPECISLTSLEDPVILSKIR QNLKEKHARHIADLRAYYESEINSLKQKLEA KEISGVEDWKITNQILVDRCGQLDSALHEATS RVRTLENKNLLEIEVNDLREFSAASSASKI LQERIEEMRTSSKEKDNTIIRLKSRLQDLEEAF ENAYKLSDDKEAQLKQENKMFQDLLGEYES LGKEHRRVKDALNTTENKLLDAYTQISDLKR MISKLEAQVKQVEHENMLSLRHNSRIHVRPS RANTLATSDVSRRKWLIPGAEYSIFTGQPLDT QDSNVDNQLEETCSLGHRSPLEKDSSP/GSSST SLLIKKQRETSDTPIMRALKELDEGKIFKNWG TQTEKEDTSNSLL*/INPRQTETSVNASRSPEK CAQQRQKRLNSASQRSSSLPPSNRKSSTPTKR EIMLTPVTVAYSPKRSPKENLSPGFSHLLSKN ESSPIREKTYSEKATDNHVNHSSCPEPVPNGV KKVSVRTAWEKNKSVSYEQCKPVSVTPQGN DFEYTAKIRTLAETERFFDELTKEKDQIEAAL SRMPSPGGRITLQTRLNQVKCLSLNLL 368 1718 A 3163 2 2350 EFKSGGCGAGLVAAGAVLVLYPASRAGERT RVPGSPAPSSLPLHSPGACGTEVDMDPQRSPL LEVKGNIELKRPLIKAPSQLPLSGSRLKRRPDQ MEDGLEPEKKRTRGLGATTKITTSHPRVPSLT TVPQTQGQTTAQKVSKKTGPRCSTAIATGLK NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP]						
VTSQLPGFPKYPSHTKASPVDSWKNQTFQNE SRTSSTFPSVYTITISNDISVNTVDEENTVMVAS ASVSQSQLPGTANSVPECISLTSLEDPVILSKIR QNLKEKHARHIADLRAYYESEINSLKQKLEA KEISGVEDWKITNQILVDRCGQLDSALHEATS RVRTLENKNNLLEIEVNDLRERFSAASSASKI LQERIEEMRTSSKEKDNTIIRLKSRLQDLEEAF ENAYKLSDDKEAQLKQENKMFQDLLGEYES LGKEHRRVKDALNTTENKLLDAYTQISDLKR MISKLEAQVKQVEHENMLSLRHNSRIHVRPS RANTLATSDVSRRKWLIPGAEYSIFTGQPLDT QDSNVDNQLEETCSLGHRSPLEKDSSP/GSSST SLLIKKQRETSDTPIMRALKELDEGKIFKNWG TQTEKEDTSNSLL*/INPRQTETSVNASRSPEK CAQQRQKRLNSAASQRSSSLPPSNRKSSTPTKR EIMLTPVTVAYSPKRSPKENLSPGFSHLLSKN ESSPIREKTYSEKATDNHVNHSSCPEPVPNGV KKVSVRTAWEKNKSVSYEQCKPVSVTPQGN DFEYTAKIRTLAETERFFDELTKEKDQIEAAL SRMPSPGGRITLQTRLNQVKCLSLNLL 368 1718 A 3163 2 2350 EFKSGGCGAGLVAAGAVLVLYPASRAGERT RVPGSPAPSSLPLHSPGACGTEVDMDPQRSPL LEVKGNIELKRPLIKAPSQLPLSGSRLKRRPDQ MEDGLEPEKKRTRGLGATTKITTSHPRVPSLT TVPQTQGQTTAQKVSKKTGPRCSTAIATGLK NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP	i						
SRTSSTFPSVYTITSNDISVNTVDEENTVMVAS ASVSQSQLPGTANSVPECISLTSLEDPVILSKIR QNLKEKHARHIADLRAYYESEINSLKQKLEA KEISGVEDWKITNQIL VDRCGQLDSALHEATS RVRTLENKNNLLEIEVNDLRERFSAASSASKI LQERIEEMRTSSKEKDNTIIRLKSRLQDLEEAF ENAYKLSDDKEAQLKQENKMPQDLLGEYES LGKEHRRVKDALNTTENKLLDAYTQISDLKR MISKLEAQVKQVEHENMLSLRHNSRIHVRPS RANTLATSDVSRRKWLIPGAEYSIFTGQPLDT QDSNVDNQLEETCSLGHRSPLEKDSSP/GSSST SLLIKKQRETSDTPIMRALKELDEGKIFKNWG TQTEKEDTSNSLL*/INPRQTETSVNASRSPEK CAQQRQKRLNSASQRSSSLPPSNRKSSTPTKR EIMLTPVTVAYSPKRSPKENLSPGFSHLLSKN ESSPIREKTYSEKATDNHVNHSSCPEPVPNGV KKVSVRTAWEKNKSVSYEQCKPVSVTPQGN DFEYTAKIRTLAETERFFDELTKEKDQIEAAL SRMPSPGGRITLQTRLNQVKCLSLNLL 368 1718 A 3163 2 2350 EFKSGGCGAGLVAAGAVLVLYPASRAGERT RVPGSPAPSSLPLHSPGACGTEVDMDPQRSPL LEVKGNIELKRPLIKAPSQLPLSGSRLKRRPDQ MEDGLEPEKRTRGLGATTKITTSHPRVPSLT TVPQTQGQTTAQKVSKKTGPRCSTAIATGLK NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP				[NALDDRISFSPDSVLEPSMSSPSDIDSFSQASN
ASVSQSQLPGTANSVPECISLTSLEDPVILSKIR QNLKEKHARHIADLRAYYESEINSLKQKLEA KEISGVEDWKITNQILVDRCGQLDSALHEATS RVRTLENKNNLLEIEVNDLRERFSAASSASKI LQERIEEMRTSSKEKDNTIIRLKSRLQDLEEAF ENAYKLSDDKEAQLKQENKMFQDLLGEYES LGKEHRRVKDALNTTENKLLDAYTQISDLKR MISKLEAQVKQVEHENMLSLRHNSRIHVRPS RANTLATSDVSRKWLIPGAEYSIFTGQPLDT QDSNVDNQLEETCSLGHRSPLEKDSSPJGSSST SLLIKKQRETSDTPIMRALKELDEGKIFKNWG TQTEKEDTSNSLL*/INPRQTETSVNASRSPEK CAQQRQKRLNSASQRSSSLPPSNRKSSTPTKR EIMLTPVTVAYSPKRSPKENLSPGFSHLLSKN ESSPIREKTYSEKATDNHVNHSSCPEPVPNGV KKVSVRTAWEKNKSVSYEQCKPVSVTPQGN DFEYTAKIRTLAETERFFDELTKEKDQIEAAL SRMPSPGGRITLQTRLNQVKCLSLNIL 368 1718 A 3163 2 2350 EFKSGGCGAGLVAAGAVLVLYPASRAGERT RVPGSPAPSSLPLHSPGACGTEVDMDPQRSPL LEVKGNIELKRPLIKAPSQLPLSGSRLKRRPDQ MEDGLEPEKKRTRGLGATTKITTSHPRVPSLT TVPQTQGQTTAQKVSKKTGPRCSTAIATGLK NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP				1 1			VTSQLPGFPKYPSHTKASPVDSWKNQTFQNE
QNLKEKHARHIADLRAYYESEINSLKQKLEA KEISGVEDWKITNQILVDRCGQLDSALHEATS RVRTLENKNNLLEIEVNDLRERFSAASSASKI LQERIEEMRTSSKEKDNTIIRLKSRLQDLEEAF ENAYKLSDDKEAQLKQENKMFQDLLGEYES LGKEHRRVKDALNTTENKLLDAYTQISDLKR MISKLEAQVKQVEHENMLSLRHNSRIHVRPS RANTLATSDVSRRKWLIPGAEYSIFTGQPLDT QDSNVDNQLEETCSLGHRSPLEKDSSP/GSSST SLLIKKQRETSDTPIMRALKELDEGKIFKNWG TQTEKEDTSNSLL*/INPRQTETSVNASRSPEK CAQQRQKRLNSASQRSSSLPPSNRKSSTPTKR EIMLTPVTVAYSPKRSPKENLSPGFSHLLSKN ESSPIREKTYSEKATDNHVNHSSCPEPVPNGV KKVSVRTAWEKNKSVSYEQCKPVSVTPQGN DFEYTAKIRTLAETERFFDELTKEKDQIEAAL SRMPSPGGRITLQTRLNQVKCLSLNLL SRMPSPGGRITLQTRLNQVKCLSLNLL SRMPSPGRITLQTRLNQVKCLSLNLL EFKSGCGAGLVAAGAVLVLYPASRAGERT RVPGSPAPSSLPLHSPGACGTEVDMDPQRSPL LEVKGNIELKRPLIKAPSQLPLSGSRLKRRPDQ MEDGLEPEKKRTRGLGATTKITTSHPRVPSLT TVPQTQGQTTAQKVSKKTGPRCSTAIATGLK NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP							SRTSSTFPSVYTITSNDISVNTVDEENTVMVAS
KEISGVEDWKITNQILVDRCGQLDSALHEATS RVRTLENKNNLLEIEVNDLRERFSAASSASKI LQERIEEMRTSSKEKDNTIIRLKSRLQDLEEAF ENAYKLSDDKEAQLKQENKMFQDLLGEYES LGKEHRRVKDALNTTENKLLDAYTQISDLKR MISKLEAQVKQVEHENMLSLRHNSRIHVRPS RANTLATSDVSRRKWLIPGAEYSIFTGQPLDT QDSNVDNQLEETCSLGHRSPLEKDSSP/GSSST SLLIKKQRETSDTPIMRALKELDEGKIFKNWG TQTEKEDTSNSLL*/INPRQTETSVNASRSPEK CAQQRQKRLNSASQRSSSLPPSNRKSSTPYKR EIMLTPVTVAYSPKRSPKENLSPGFSHLLSKN ESSPIREKTYSEKATDNHVNHSSCPEPVPNGV KKVSVRTAWEKNKSVSYEQCKPVSVTPQGN DFEYTAKIRTLAETERFFDELTKEKDQIEAAL SRMPSPGGRITLQTRLNQVKCLSLNLL 368 1718 A 3163 2 2350 EFKSGGCGAGLVAAGAVLVLYPASRAGERT RVPGSPAPSSLPLHSPGACGTEVDMDPQRSPL LEVKGNIELKRPLKAPSQLPLSGSRLKRRPDQ MEDGLEPEKKRTRGLGATTKITTSHPRVPSLT TVPQTQGQTTAQKVSKKTGPRCSTAIATGLK NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP							ASVSQSQLPGTANSVPECISLTSLEDPVILSKIR
KEISGVEDWKITNQILVDRCGQLDSALHEATS RVRTLENKNNLLEIEVNDLRERFSAASSASKI LQERIEEMRTSSKEKDNTIIRLKSRLQDLEEAF ENAYKLSDDKEAQLKQENKMFQDLLGEYES LGKEHRRVKDALNTTENKLLDAYTQISDLKR MISKLEAQVKQVEHENMLSLRHNSRIHVRPS RANTLATSDVSRRKWLIPGAEYSIFTGQPLDT QDSNVDNQLEETCSLGHRSPLEKDSSP/GSSST SLLIKKQRETSDTPIMRALKELDEGKIFKNWG TQTEKEDTSNSLL*/INPRQTETSVNASRSPEK CAQQRQKRLNSASQRSSSLPPSNRKSSTPYKR EIMLTPVTVAYSPKRSPKENLSPGFSHLLSKN ESSPIREKTYSEKATDNHVNHSSCPEPVPNGV KKVSVRTAWEKNKSVSYEQCKPVSVTPQGN DFEYTAKIRTLAETERFFDELTKEKDQIEAAL SRMPSPGGRITLQTRLNQVKCLSLNLL 368 1718 A 3163 2 2350 EFKSGGCGAGLVAAGAVLVLYPASRAGERT RVPGSPAPSSLPLHSPGACGTEVDMDPQRSPL LEVKGNIELKRPLKAPSQLPLSGSRLKRRPDQ MEDGLEPEKKRTRGLGATTKITTSHPRVPSLT TVPQTQGQTTAQKVSKKTGPRCSTAIATGLK NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP							QNLKEKHARHIADLRAYYESEINSLKOKLEA
RVRTLENKNNLLEIEVNDLRERFSAASSASKI LQERIEEMRTSSKEKDNTIIRLKSRLQDLEEAF ENAYKLSDDKEAQLKQENKMFQDLLGEYES LGKEHRRVKDALNTTENKLLDAYTQISDLKR MISKLEAQVKQVEHENMLSLRHNSRIHVRPS RANTLATSDVSRRKWLIPGAEYSIFTGQPLDT QDSNVDNQLEETCSLGHRSPLEKDSSP/GSSST SLLIKKQRETSDTPIMRALKELDEGKIFKNWG TQTEKEDTSNSLL*/INPRQTETSVNASRSPEK CAQQRQKRLINSASQRSSSLPPSNRKSSTPTKR EIMLTPVTVAYSPKRSPKENLSPGFSHLLSKN ESSPIREKTYSEKATDNHVNHSSCPEPVPNGV KKVSVRTAWEKNKSVSYEQCKPVSVTPQGN DFEYTAKIRTLAETERFFDELTKEKDQIEAAL SRMPSPGGRITLQTRLNQVKCLSLNLL 368 1718 A 3163 2 2350 EFKSGGCGAGLVAAGAVLVLYPASRAGERT RVPGSPAPSSLPLHSPGACGTEVDMDPQRSPL LEVKGNIELKRPLIKAPSQLPLSGSRLKRRPDQ MEDGLEPEKKTRGLGATTKITTSHPRVPSLT TVPQTQGQTTAQKVSKKTGPRCSTAIATGLK NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP				j	ļ	ļ	
LQERIEEMRTSSKEKDNTIIRLKSRLQDLEEAF ENAYKLSDDKEAQLKQENKMFQDLLGEYES LGKEHRRVKDALNTTENKLLDA YTQISDLKR MISKLEAQVKQVEHENMLSLRHINSRIHVRPS RANTLATSDVSRRKWLIPGAEYSIFTGPPLDT QDSNVDNQLEETCSLGHRSPLEKDSSP/GSSST SLLIKKQRETSDTPIMRALKELDEGKIFKNWG TQTEKEDTSNSLL*/INPRQTETSVNASRSPEK CAQQRQKRLNSASQRSSSLPPSNRKSSTPTKR EIMLTPVTVAYSPKRSPKENLSPGFSHLLSKN ESSPIREKTYSEKATDNHVNHSSCPEPVPNGV KKVSVRTAWEKNKSVSYEQCKPVSVTPQGN DFEYTAKIRTLAETERFFDELTKEKDQIEAAL SRMPSPGGRITLQTRLNQVKCLSLNLL 368 1718 A 3163 2 2350 EFKSGGCGAGLVAAGAVLVLYPASRAGERT RVPGSPAPSSLPLHSPGACGTEVDMDPQRSPL LEVKGNIELKRPLIKAPSQLPLSGSRLKRRPDQ MEDGLEPEKKRTRGLGATTKITTSHPRVPSLT TVPQTQGQTTAQKVSKKTOPRCSTAIATGLK NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP					:		
ENAYKLSDDKEAQLKQENKMFQDLLGEYES LGKEHRRVKDALNTTENKLLDAYTQISDLKR MISKLEAQVKQVEHENMLSLRHNSRIHVRPS RANTLATSDVSRRKWLIPGAEYSIFTGQPLDT QDSNVDNQLEETCSLGHRSPLEKDSSP/GSSST SLLIKKQRETSDTPIMRALKELDEGKIFKNWG TQTEKEDTSNSLL*/INPRQTETSVNASRSPEK CAQQRQKRLNSASQRSSSLPPSNRKSSTTPTKR EIMLTPVTVAYSPKRSPKENLSPGFSHLLSKN ESSPIREKTYSEKATDNHVNHSSCPEPVPNGV KKVSVRTAWEKNKSVSYEQCKPVSVTPQGN DFEYTAKIRTLAETERFFDELTKEKDQIEAAL SRMPSPGGRITLQTRLNQVKCLSLNLL 368 1718 A 3163 2 2350 EFKSGGCGAGLVAAGAVLVLYPASRAGERT RVPGSPAPSSLPLHSPGACGTEVDMDPQRSPL LEVKGNIELKRPLIKAPSQLPLSGSRLKRRPDQ MEDGLEPEKKRTRGLGATTKITTSHPRVPSLT TVPQTQGQTTAQKVSKKTOPRCSTAIATGLK NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP							
LGKEHRRVKDALNTTENKLLDAYTQISDLKR MISKLEAQVKQVEHENMLSLRHNSRIHVRPS RANTLATSDVSRRKWLIPGAEYSIFTGQPLDT QDSNVDNQLEETCSLGHRSPLEKDSSP/GSSST SLLIKKQRETSDTPIMRALKELDEGKIFKNWG TQTEKEDTSNSLL*/INPRQTETSVNASRSPEK CAQQRQKRLNSASQRSSSLPPSNRKSSTPTKR EIMLTPVTVAYSPKRSPKENLSPGFSHLLSKN ESSPIREKTYSEKATDNHVNHSSCPEPVPNGV KKVSVRTAWEKNKSVSYEQCKPVSVTPQGN DFEYTAKIRTLAETERFFDELTKEKDQIEAAL SRMPSPGGRITLQTRLNQVKCLSLNLL 368 1718 A 3163 2 2350 EFKSGGCGAGLVAAGAVLVLYPASRAGERT RVPGSPAPSSLPLHSPGACGTEVDMDPQRSPL LEVKGNIELKRPLIKAPSQLPLSGSRLKRRPDQ MEDGLEPEKKRTRGLGATTKITTSHPRVPSLT TVPQTQGQTTAQKVSKKTOPRCSTAIATGLK NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP	!				l		
MISKLEAQVKQVEHENMLSLRHNSRIHVRPS RANTLATSDVSRRKWLIPGAEYSIFTGQPLDT QDSNVDNQLEETCSLGHRSPLEKDSSP/GSSST SLLIKKQRETSDTPIMRALKELDEGKIFKNWG TQTEKEDTSNSLL*/INPRQTETSVNASRSPEK CAQQRQKRLNSASQRSSSLPPSNRKSSTPTKR EIMLTPVTVAYSPKRSPKENLSPGFSHLLSKN ESSPIREKTYSEKATDNHVNHSSCPEPVPNGV KKVSVRTAWEKNKSVSYEQCKPVSVTPQGN DFEYTAKIRTLAETERFFDELTKEKDQIEAAL SRMPSPGGRITLQTRLNQVKCLSLNLL 368 1718 A 3163 2 2350 EFKSGGCGAGLVAAGAVLVLYPASRAGERT RVPGSPAPSSLPLHSPGACGTEVDMDPQRSPL LEVKGNIELKRPLIKAPSQLPLSGSRLKRRPDQ MEDGLEPEKKRTRGLGATTKITTSHPRVPSLT TVPQTQGQTTAQKVSKKTOPRCSTAIATGLK NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP							
RANTLATSDVSRRKWLIPGAEYSIFTGQPLDT QDSNVDNQLEETCSLGHRSPLEKDSSP/GSSST SLLIKKQRETSDTPIMRALKELDEGKIFKNWG TQTEKEDTSNSLL*/INPRQTETSVNASRSPEK CAQQRQKRLNSASQRSSSLPPSNRKSSTPTKR EIMLTPVTVAYSPKRSPKENLSPGFSHLLSKN ESSPIREKTYSEKATDNHVNHSSCPEPVPNGV KKVSVRTAWEKNKSVSYEQCKPVSVTPQGN DFEYTAKIRTLAETERFFDELTKEKDQIEAAL SRMPSPGGRITLQTRLNQVKCLSLNLL 368 1718 A 3163 2 - 2350 EFKSGGCGAGLVAAGAVLVLYPASRAGERT RVPGSPAPSSLPLHSPGACGTEVDMDPQRSPL LEVKGNIELKRPLIKAPSQLPLSGSRLKRRPDQ MEDGLEPEKKRTRGLGATTKITTSHPRVPSLT TVPQTQGQTTAQKVSKKTOPRCSTAIATGLK NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP					-		
QDSNVDNQLEETCSLGHRSPLEKDSSP/GSSST SLLIKKQRETSDTPIMRALKELDEGKIFKNWG TQTEKEDTSNSLL*/INPRQTETSVNASRSPEK CAQQRQKRLNSASQRSSSLPPSNRKSSTPTKR EIMLTPVTVAYSPKRSPKENLSPGFSHLLSKN ESSPIREKTYSEKATDNHVNHSSCPEPVPNGV KKVSVRTAWEKNKSVSYEQCKPVSVTPQGN DFEYTAKIRTLAETERFFDELTKEKDQIEAAL SRMPSPGGRITLQTRLNQVKCLSLNLL 368 1718 A 3163 2 - 2350 EFKSGGCGAGLVAAGAVLVLYPASRAGERT RVPGSPAPSSLPLHSPGACGTEVDMDPQRSPL LEVKGNIELKRPLIKAPSQLPLSGSRLKRRPDQ MEDGLEPEKKRTRGLGATTKITTSHPRVPSLT TVPQTQGQTTAQKVSKKTGPRCSTAIATGLK NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP					}		
SLLIKKQRETSDTPIMRALKELDEGKIFKNWG TQTEKEDTSNSLL*/INPRQTETSVNASRSPEK CAQQRQKRLNSASQRSSSLPPSNRKSSTPTKR EIMLTPVTVAYSPKRSPKENLSPGFSHLLSKN ESSPIREKTYSEKATDNHVNHSSCPEPVPNGV KKVSVRTAWEKNKSVSYEQCKPVSVTPQGN DFEYTAKIRTLAETERFFDELTKEKDQIEAAL SRMPSPGGRITLQTRLNQVKCLSLNLL 368 1718 A 3163 2 2350 EFKSGGCGAGLVAAGAVLVLYPASRAGERT RVPGSPAPSSLPLHSPGACGTEVDMDPQRSPL LEVKGNIELKRPLIKAPSQLPLSGSRLKRRPDQ MEDGLEPEKKRTRGLGATTKITTSHPRVPSLT TVPQTQGQTTAQKVSKKTGPRCSTAIATGLK NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP							
TQTEKEDTSNSLL*/INPRQTETSVNASRSPEK CAQQRQKRLNSASQRSSSLPPSNRKSSTPTKR EIMLTPVTVAYSPKRSPKENLSPGFSHLLSKN ESSPIREKTYSEKATDNHVNHSSCPEPVPNGV KKVSVRTAWEKNKSVSYEQCKPVSVTPQGN DFEYTAKIRTLAETERFFDELTKEKDQIEAAL SRMPSPGGRITLQTRLNQVKCLSLNLL 368 1718 A 3163 2 - 2350 EFKSGGCGAGLVAAGAVLVLYPASRAGERT RVPGSPAPSSLPLHSPGACGTEVDMDPQRSPL LEVKGNIELKRPLIKAPSQLPLSGSRLKRRPDQ MEDGLEPEKKRTRGLGATTKITTSHPRVPSLT TVPQTQGQTTAQKVSKKTGPRCSTAIATGLK NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP		,			1		
CAQQRQKRLNSASQRSSSLPPSNRKSSTPTKR EIMLTPVTVAYSPKRSPKENLSPGFSHLLSKN ESSPIREKTYSEKATDNHVNHSSCPEPVPNGV KKVSVRTAWEKNKSVSYEQCKPVSVTPQGN DFEYTAKIRTLAETERFFDELTKEKDQIEAAL SRMPSPGGRITLQTRLNQVKCLSLNLL 368 1718 A 3163 2 - 2350 EFKSGGCGAGLVAAGAVLVLYPASRAGERT RVPGSPAPSSLPLHSPGACGTEVDMDPQRSPL LEVKGNIELKRPLIKAPSQLPLSGSRLKRRPDQ MEDGLEPEKRTTRGLGATTKITTSHPRVPSLT TVPQTQGQTTAQKVSKKTGPRCSTAIATGLK NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP							SLLIKKQKETSDTPIMRALKELDEGKIFKNWG
EIMLTPVTVAYSPKRSPKENLSPGFSHLLSKN ESSPIREKTYSEKATDNHVNHSSCPEPVPNGV KKVSVRTAWEKNKSVSYEQCKPVSVTPQGN DFEYTAKIRTLAETERFFDELTKEKDQIEAAL SRMPSPGGRITLQTRLNQVKCLSLNLL 368 1718 A 3163 2 2350 EFKSGGCGAGLVAAGAVLVLYPASRAGERT RVPGSPAPSSLPLHSPGACGTEVDMDPQRSPL LEVKGNIELKRPLIKAPSQLPLSGSRLKRRPDQ MEDGLEPEKRTTRGLGATTKITTSHPRVPSLT TVPQTQGQTTAQKVSKKTOPRCSTAIATGLK NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP				İ	Į.		
ESSPIREKTYSEKATDNHVNHSSCPEPVPNGV KKVSVRTAWEKNKSVSYEQCKPVSVTPQGN DFEYTAKIRTLAETERFFDELTKEKDQIEAAL SRMPSPGGRITLQTRLNQVKCLSLNLL 368 1718 A 3163 2 - 2350 EFKSGGCGAGLVAAGAVLVLYPASRAGERT RVPGSPAPSSLPLHSPGACGTEVDMDPQRSPL LEVKGNIELKRPLIKAPSQLPLSGSRLKRRPDQ MEDGLEPEKKRTRGLGATTKITTSHPRVPSLT TVPQTQGQTTAQKVSKKTOPRCSTAIATGLK NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP							
KKVSVRTAWEKNKSVSYEQCKPVSVTPQGN DFEYTAKIRTLAETERFFDELTKEKDQIEAAL SRMPSPGGRITLQTRLNQVKCLSLNLL 368 1718 A 3163 2 - 2350 EFKSGGCGAGLVAAGAVLVLYPASRAGERT RVPGSPAPSSLPLHSPGACGTEVDMDPQRSPL LEVKGNIELKRPLIKAPSQLPLSGSRLKRRPDQ MEDGLEPEKKRTRGLGATTKITTSHPRVPSLT TVPQTQGQTTAQKVSKKTOPRCSTAIATGLK NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP							EIMLTPVTVAYSPKRSPKENLSPGFSHLLSKN
DFEYTAKIRTLAETERFFDELTKEKDQIEAAL SRMPSPGGRITLQTRLNQVKCLSLNLL 368 1718 A 3163 2 2350 EFKSGGCGAGLVAAGAVLVLYPASRAGERT RVPGSPAPSSLPLHSPGACGTEVDMDPQRSPL LEVKGNIELKRPLIKAPSQLPLSGSRLKRRPDQ MEDGLEPEKKRTRGLGATTKITTSHPRVPSLT TVPQTQGQTTAQKVSKKTGPRCSTAIATGLK NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP		}					ESSPIREKTYSEKATDNHVNHSSCPEPVPNGV
DFEYTAKIRTLAETERFFDELTKEKDQIEAAL SRMPSPGGRITLQTRLNQVKCLSLNLL 368 1718 A 3163 2 2350 EFKSGGCGAGLVAAGAVLVLYPASRAGERT RVPGSPAPSSLPLHSPGACGTEVDMDPQRSPL LEVKGNIELKRPLIKAPSQLPLSGSRLKRRPDQ MEDGLEPEKKRTRGLGATTKITTSHPRVPSLT TVPQTQGQTTAQKVSKKTGPRCSTAIATGLK NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP		ļ					
SRMPSPGGRITLQTRLNQVKCLSLNLL 368 1718 A 3163 2 - 2350 EFKSGGCGAGLVAAGAVLVLYPASRAGERT RVPGSPAPSSLPLHSPGACGTEVDMDPQRSPL LEVKGNIELKRPLIKAPSQLPLSGSRLKRRPDQ MEDGLEPEKKRTRGLGATTKITTSHPRVPSLT TVPQTQGQTTAQKVSKKTGPRCSTAIATGLK NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP			[[[
368 1718 A 3163 2 2350 EFKSGGCGAGLVAAGAVLVLYPASRAGERT RVPGSPAPSSLPLHSPGACGTEVDMDPQRSPL LEVKGNIELKRPLIKAPSQLPLSGSRLKRRPDQ MEDGLEPEKKRTRGLGATTKITTSHPRVPSLT TVPQTQGQTTAQKVSKKTGPRCSTAIATGLK NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP					1		
RVPGSPAPSSLPLHSPGACGTEVDMDPQRSPL LEVKGNIELKRPLIKAPSQLPLSGSRLKRRPDQ MEDGLEPEKKRTRGLGATTKITTSHPRVPSLT TVPQTQGQTTAQKVSKKTGPRCSTAIATGLK NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP	368	1718	A	3163	2 -	2350	
LEVKGNIELKRPLIKAPSQLPLSGSRLKRRPDQ MEDGLEPEKKRTRGLGATTKITTSHPRVPSLT TVPQTQGQTTAQKVSKKTGPRCSTAIATGLK NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP				3.05	-	-550	
MEDGLEPEKKRTRGLGATTKITTSHPRVPSLT TVPQTQGQTTAQKVSKKTGPRCSTAIATGLK NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP							
TVPQTQGQTTAQKVSKKTGPRCSTAIATGLK NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP			ĺ	l			
NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP	}			}]	
NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP AWDLKGQLCDLNAELKRCRERTQTLDQENQ]		[{	[[
AWDLKGQLCDLNAELKRCRERTQTLDQENQ	1				1		NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP
· · · · · · · · · · · · · · · · · · ·							AWDLKGQLCDLNAELKRCRERTQTLDQENQ

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				·		QLQDQLRDAQQQVKALGTERTTLEGHLAKV QAQAEQGQQELKNLRACVLELEERLSTQEGL VQELQKKQVELQEERRGLMSQLEEKERRLQT SEAALSSSQAEVASLRQETVAQAALLTEREER LHGLEMERRLHNQLQELKGNIRVFCRVRPV LPGEPTPPPGLLLFPSGPGGPSDPPTRLSLSRSD ERRGTLSGAPAPPTRHDFSFDRVFPPGSQQDE VFEEIAMLVQSALDGYPVCIFAYGQTGSGKTF TMEGGPGGDPQLEGLIPRALRHLFSVAQELSG QGWTYSFVASYVEIYNETVRDLLATGTRKGQ GGECEIRRAGPGSEELTVTNARYVPVSCEKEV DALLHLARQNRAVARTAQNERSSRSHSVFQL QISGEHSSRGLQCGAPLSLVDLAGSERLDPGL ALGPGERERLRETQAINSSLSTLGLVIMALSN KESHVPYRNSKLTYLLQNSLGGSAKMLMFV NISPLEENVSESLNSLRFASKVEPSVLFGTAQS NRKWKTDPDLCVCVCVCVCVCVCVCVCVP MSMYRVRGGRVAGGFIGWRAPCPRAIK
369	1719	A -	3165	365	12	GYTSQGRWIDIERGPLTANTESLHENNFNALP GYIRKIE*I*IYKKN*INFGGVGLLNIVKISILS/K IYRFDAIPVKILTRFFINLDKLILKFVLKTKIAK NRIKTFYIMRRKKLGDSS
370	1720	Α.	3170	393	42	GASISPSAVIDGVEGLKPMQEQEAQEAGPCLD *HMAPEQWVAPR\RLLFRLIFSVLHALIIAAAA QSSAEEDEDPRN*GQSSEDQAPNQNGLIVIVH RVHVPLGAAATVPVHRSHFPR
371	1721	A	3173	770	510	GNGGCGLSQIPPSHLGAFSRGSLLSRG\DPRGP PPHPVIFFVFVVE\QGFTVLARMVSIS*PCDPP ALASQSAGITGVSHLARPQNLYF
372	1722	A	3180	381	76	RVLHHDNVPAHSSPQKREISOEFQLEIRHLP*S PDLAPSGCFLFLNLKNIFK\GTHFSLVDNVKK TVSTWLH/SQNAQFYKDRLNGWYHCLQKCL
	1723	A	3181	410	14101	RREVAGPEGKGLLLASAHTMLTPPLLLLPLL SALVAAAIDAPKTCSPKQFACRDQITCISKGW RCDGERDCPDGSDEAPEICPQSKAQRCQPNE HNCLGTELCVPMSRLCNGVQDCMDGSDEGP HCRELQGNCSRLGCQHHCVPTLDGPTCYCNS SFQLQADGKTCKDFDECSVYGTCSQLCTNTD GSFICGCVEGYLLQPDNRSCKAKNEPVDRPP VLIANSQNILATYLSGAQVSTITPTSTRQTTA MDFSYANETVCWVHVGDSAAQTQLKCARM PGLKGFVDEHTINISLSLHHVEQMAIDWLTGN FYFVDDIDDRIFVCNRNGDTCVTLLDLELYNP KGIALDPAMGKVFFTDYGQIPKVERCDMDG QNRTKLVDSKIVFPHGITLDLVSRLVYWADA YLDYIEVVDYEGKGRQTIIQGILIEHLYGLTVF ENYLYATNSDNANAQQKTSVIRVNRFNSTEY QVVTRVDKGGALHIYHQRRQPRVRSHACEN DQYGKPGGCSDICLLANSHKARTCRCRSGFS LGSDGKSCKKPEHELFLVYGKGRPGIIRGMD MGAKVPDEHMIPIENLMNPRALDFHAETGFI YFADTTSYLIGRQKIDGTERETILKDGIHNVE GVAVDWMGDNLYWTDDGPKKTISVARLEK AAQTRKTLIEGKMTHPRAIVVDPLNGWMYW TDWEEDPKDSRRGRLERAWMDGSHRDIFVT SKTVLWPNGLSLDIPAGRLYWVDAFYDRIETI LLNGTDRKIVYEGPELNHAFGLCHHGNYLFW TEYRSGSVYRLERGVGGAPPTVTLLRSENPPI

SEQ ID NO: of NO: of nucl- cotide seq- uence uence SEQ ID NO: of NO: of nucl- cotide seq- uence SEQ ID NO: of NO: of nucl- cotide seq- uence SEQ ID NO: of NO: of nucl- in USSN location corresponding corresponding to last amino corresponding to first amino acid sequence (A predicted end nucleotide location corresponding to last amino acid sequence (A predicted end nucleotide location pucleotide location corresponding to last amino acid sequence (A predicted end nucleotide location pucleotide location corresponding to last amino acid sequence (A predicted end nucleotide location pucleotide location corresponding to last amino acid sequence (A predicted end nucleotide location pucleotide location corresponding to last amino acid sequence (A predicted end nucleotide location pucleotide fresidue of period predicted end nucleotide location pucleotide fresidue of predicted end nucleotide location pucleotide fresidue of predicted end nucleotide fresponding last amino acid sequence (A predicted end nucleotide location pucleotide fresponding last amino acid sequence (A predicted end nucleotide location pucleotide location corresponding to last amino acid sequence (A predicted end nucleotide location pucleotide location corresponding to last amino acid sequence (A predicted end nucleotide location pucleotide fresponding pucleotide location pucleotide location pucleotide fresponding pucleotide fresponding pucleotide fresponding pucleotide fresponding pucleotide fresponding pucleotide fresponding pucleotide fresponding pucleotide fresponding pucleotide fresponding fresponding pucleotide fresponding pucleotide fresponding pucleotide fresponding pucleotide fresponding fresponding fresponding pucleotide fresponding fr	utamic Acid, ycine, H=Histidine, , L=Leucine, aragine, P=Proline, ine, S=Serine, c, W=Tryptophan,
nucleotide cotide sequence uence uence uence sequence sequence uence uence uence uence sequen	ycine, H=Histidine, , L=Leucine, aragine, P=Proline, ine, S=Serine, c, W=Tryptophan,
eotide sequence USSN location corresponding to last amino memory of peptide Sequence 914 location corresponding to last amino memory of peptide Sequence 914 location corresponding to last amino memory delication corresponding to last amino acid memory delication corresponding to last amino memory delication corresponding to last amino memory delication corresponding to last amino memory delication corresponding to last amino memory delication corresponding to last amino memory delication corresponding to last amino memory delication corresponding to last amino memory delication corresponding to last amino memory delication corresponding to last amino memory delication corresponding to last amino memory delication corresponding to last amino memory delication corresponding to last amino memory delication corresponding to last amino memory delication corresponding to last amino memory delication corresponding to last amino memory	, L=Leucine, aragine, P=Proline, ine, S=Serine, c, W=Tryptophan,
seq- uence 09/496 correspondi to last amino M=Methionine, N=Aspi ng to first acid residue Q=Glutamine, R=Argin amino acid of peptide T=Threonine, V=Valine	aragine, P=Proline, ine, S=Serine, c, W=Tryptophan,
uence 914 ng to first acid residue Q=Glutamine, R=Argin amino acid of peptide T=Threonine, V=Valine	ine, S=Serine, c, W=Tryptophan,
amino acid of peptide T=Threonine, V=Valine	, W=Tryptophan,
peptide /=possible nucleotide de	
sequence nucleotide insertion	eletion, =possible
	SNKCRVNNAGCSSLCL
ATROSPOCACA EDIO	VLDADGVTCLANPSYVP
	.CIQERWKCDGDNDCLD
	PSDRFKCENNRCIPNRW
	SNATCSARTCPPNQFSC
	DDCGDRSDESASCAYPT
	NINWRCDNDNDCGDNS
DEAGCSHSCSSTOFK	CNSGRCIPEHWTCDGD
NDCGDYSDETHANC	TNQATRPPGGCHTDEF
	DGDTDCMDSSDEKSCE
	KDSARCISKAWVCDGD
	ACRPPSHPCANNTSVC
	DGSDEGELCDQCSLNN
	VCSCPLGMELGPDNHT
CQIQSYCAKHLKCSQ	KCDQNKFSVKCSCYEG
WVLEPDGESCRSLDP	PFKPFIIFSNRHEIRRIDLH
	ALDFHLSQSALYWTDV
	LTSFEVVIQYGLATPEG
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	SNLDQIEVAKLDGTLRT
	PRDGILFWTDWDASLP
	HRETGSGGWPNGLTV
	AIYSARYDGSGHMEVL
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	GGEVYWTDWRTNTLA
	QRTNTQPFDLQVYHPSR
	GPCSHLCLINYNRTVSC CYEFKKFLLYAROMEIR
	CYPDIDNYTVLDYDARE
	RAFINGTGVETVVSADL
	LFWTSYDTNKKQINVA
	QPHGLVVHPLRGKLY
	SNRTLLFSGQKGPVGL
	HTINRCNLDGSGLEVID
AMRSQLGKATALAIN	MGDKLWWADQVSEKM
	NSTTLVMHMKVYDESI
QLDHKGTNPCSVNNC	GDCSQLCLPTSETTRSC
	CEGVGSFLLYSVHEGIR
	VSGTSLAVGIDFHAEND
TIYWVDMGLSTISRAI	KRDQTWREDVVTNGIG
RVEGIAVDWIAGNIY	WTDQGFDVIEVARLNG
STRY VVISQGLDKPRA	AITVHPEKGYLFWTEW
DVODOVI VIVODANT	RVVLVNVSISWPNGISV
	DKIERIDLETGENREVV DFIYWSDRTHANGSIK
	GIGVQLKDIKVFNRDR
	QQLCLYRGRGQRACA
	EYAGYLLYSERTILKSI
	EDPEHMKNVIALAFDY
	IHFGNIQQINDDGSRRIT
	GWDTLYWTSYTTSTIT
	TVITMSGDDHPRAFVL
	QHPSIMRAALSGANVL
	iraeklyfsdatldkie
RCEYDGSHRYVILKSE	EPVHPFGLAVYGEHIF
	CHVGSNMKLLRVDIPQ
QPMGIIAVANDTNSCE	ELSPCRINNGGCQDLCL
	ULQDDLTCRAVNSSCR
	SLTCDGVPHCKDKSDE
	QCSNGRCVSNMLWCN
	KTACGVGEFRCRDGTC
IGNSSRCNQFVDCEDA	ASDEMNCSATDCSSYF

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl- eotide seq-	peptide seq- uence		in USSN 09/496 914	nucleotide location correspondi	location corresponding to last amino	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
uence			714	ng to first amino acid residue of	acid residue of peptide sequence	Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon.
				peptide sequence	soquenes	/=possible nucleotide deletion, \=possible nucleotide insertion
						RLGVKGVLFQPCERTSLCYAPSWVCDGAND CGDYSDERDCPGVKRPRCPLNYFACPSGRCIP
						MSWTCDKEDDCEHGEDETHCNKFCSEAQFE CQNHRCISKQWLCDGSDDCGDGSDEAAHCE
					· -	GKTCGPSSFSCPGTHVCVPERWLCDGDKDCA DGADESIAAGCLYNSTCDDREFMCQNRQCIP
						KHFVCDHDRDCADGSDESPECEYPTCGPSEF RCANGRCLSSRQWECDGENDCHDQSDEAPK NPHCTSPEHKCNASSQFLCSSGRCVAEALLCN
						GQDDCGDSSDERGCHINECLSRKLSGCSQDC EDLKIGFKCRCRPGFRLKDDGRTCADVDECS
					:	TTFPCSQRCINTHGSYKCLCVEGYAPRGGDP HSCKAVTDEEPFLIFANRYYLRKLNLDGSNY
						TLLKQGLNNAVALDFDYREQMIYWTDVTTQ GSMIRRMHLNGSNVQVLHRTGLSNPDGLAV
						DWVGGNLYWCDKGRDTIEVSKLNGAYRTVL VSSGLREPRALVVDVQNGYLYWTDWGDHSL
						IGRIGMDGSSRSVIVDTKITWPNGLTLDYVTE RIYWADAREDYIEFASLDGSNRHVVLSQDIPH IFALTLFEDYVYWTDWETKSINRAHKTTGTN
						KTLLISTLHRPMDLHVFHALRQPDVPNHPCK VNNGGCSNLCLLSPGGGHKCACPTNFYLGSD
						GRTCVSNCTASQFVCKNDKCIPFWWKCDTE DDCGDHSDEPPDCPEFKCRPGQFQCSTGICTN
						PAFICDGDNDCQDNSDEANCDIHVCLPSQFK CTNTNRCIPGIFRCNGQDNCGDGEDERDCPE VTCAPNQFQCSITKRCIPRVWVCDRDNDCVD
						GSDEPANCTQMTCGVDEFRCKDSGRCIPARW KCDGEDDCGDGSDEPKEECDERTCEPYQFRC
	i					KNNRCVPGRWQCDYDNDCGDNSDEESCTPR PCSESEFSCANGRCIAGRWKCDGDHDCADGS
		-				DEKDCTPRCDMDQFQCKSGHCIPLRWRCDA DADCMDGSDEEACGTGVRTCPLDEFQCNNT
				`		LCKPLAWKCDGEDDCGDNSDENPEECARFV CPPNRPFRCKNDRVCLWIGRQCDGTDNCGD GTDEEDCEPPTAHTTHCKDKKEFLCRNQRCL
						SSSLRCNMFDDCGDGSDEEDCSIDPKLTSCAT NASICGDEARCVRTEKAAYCACRSGFHTVPG
						QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS
						LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR
						RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT
						LRETLYQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV
						FEDYIYGVTYINNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL
						CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC
						QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT
						VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS
	ļ					CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL
						VAGVVFWYKRRVQGAKGFQHQRMTNGAM NVEIGNPTYKMYEGGEPDDVGGLLDADFAL

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in USSN	nucleotide	location	F-Phenylalanine, G-Glycine, H-Histidine,
eotide seq-	seq- uence		09/496	location correspondi	corresponding to last amino	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine.
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		ł		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1		l		peptide	}	/=possible nucleotide deletion, \=possible
	-	 	 	sequence		nucleotide insertion
					1	DPDKPTNFTNPVYATLYMGGHGSRHSLASTD EKRELLGRGPEDEIGDPLA
374	1724	A	3187	191	1815	CLELASAGKIPEESKALSLLAPAPTMTSLMPG
	'					AGLLPIPTPNPLTTLGVSLSSLGAIPAAALDPNI
						ATLGEIPQPPLMGNVDPSKIDEIRRTVYVGNL
	1	i	ł		ł	NSQTTTADQLLEFFKQVGEVKFVRMAGDET
!				İ		QPTRFAFVEFADQNSVPRALAFNGVMFGDRP LKINHSNNAIVKPPEMTPQAAAKELEEVMKR
	1		ļ			VREAQSFISAAIEPGWLHSTSLCNDFLGCF*RR
		1	l		í	RMYRE*APCTICGTFHLCLIINWDL*LF*AYTA
1	1	1				K*FFPPRVWKEQ*KKRR\RSRSHTRSKSRSSSK
	1	1			! !	SHSRRKRSQSKHRSRSHNRSRSRQKDRRRSK
1	1	ì				SPHKKRSKSRERRKSRSRSHSRDKRKDTREKI
						KEKERVKEKDREKEREREKEREKERGKN KDRDKEREKDREKDKEKDREREREKEHEKD
		ļ				RDKEKEKEQDKEKEREKDRSKEIDEKRKKDK
	ļ	1	1		ļ	KSRTPPRSYNASRRSRSSSRERRRRRSRSSSRS
						PRTSKTIKRKSSRSPSPRSRNKKDKKREKERD
						HISERRERERSTSMRKSSNDRDGKEKLEKNST
375	1725	A	3192	415	101	S
373	1723	A	3192	415	101	AHSSHQTRAILQEFQWDIIRHPPL\SPNLALSG F\FPNLKKSLRGTHFSSVKK\TTLTWLNSQDP
		ŀ				WF/FFYP*SPDLQIPSSFRNGLNDWYHHSOKC
		ŀ				PDLDGAYVKK
376	1726	A	3199	931	418	GV*WCDLGSPQPPPPGFKQFCLGRSSSWDYR
						HVPPHPANFVFLLETGFLHAGQAGL\GDPPAS
						ASQSAGITGVSHTWPKNHLIFYACLVIRSKRI
377	1727	A	3201	274	1285	K KTGYTSRGSPLSPQSSIDSELSTSELEDDSISM
"	*	**	3201	2,4	1205	GYKLQDLTDVQIMARLQEESLRQDYASTSAS
						VSRHSSSVSLSSGKKGTCSDQEYDQYSLEDEE
					•	EFDHLPPPQPRLPRCSPFQRGIPHSQTFSSIREC
			1			RRSPSSQYFPSNNYQQQQYYSPQAQTPDQQP
				:		NRTNGDK/PPKKYA*PSPDAKYNCH**QH\SSP
						VTVRNSQSFDSSLHGAGNGISRIQSCIPSPGQL QHRVHSVGHFPVSIRQPLKATAYVSPTVQGSS
	}		i 1			NMPLSNGLQLYSNTGIPTPNKAAASGIMGRS
					•	ALPRPSLAINGSNLPRSKIAQPVRSFLQPPKPL
						SSLSTLRDGNWRDGCY
378	1728	A	3202	112	1789	VPGVTESRPSVLRGDHLFALLSSETHQEDPIT
	[' I	YKGFVHKV\ELDRVKLSFSMSLLSRFVGWG*
				ļ		PFKVNFY/TFNRQPLRV\QHRALELTGRWLLW PMLFP\VAPRDVPLLPSDVKLKLYDRSLESNP
						EQLQAMRHIVTGTTRPAPYIIFGPPGTGKTVT
1			}			LVEAIKQVVKHLPKAHILACAPSNSGADLLC
						QRLRVHLPSSIYRLLAPSRDIRMVPEDIKPCCN
						WDAKKGEYVFPAKKKLQEYRVLITTLITAGR
			}			LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG
					İ	LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL TQKHGLGYSLLERLLTYNSLYKKGPDGYDPO
						FITKLLRNYRSHPTILDIPNOLYYEGELOACA
1					}	DVVDRERFCRWAG\LPRQGFPIIFHGVMGKD
						EREGNSPSFFNPEEAATVTSYLKLLLAPSSKK
1					1	GKARLSPRSVGVISPYRKQVEKIRYCITKLDR
}						ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP
						ETSSSFHSSPRPRPTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS
379	1729	A	3206	432	130	PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK
L		_				*LSTREAXDSXPGRQIAXXRQGGKVETTTAL

SEQ ID NO: of nucl- eotide scq- uence	SEQ ID NO: of pcptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion XKQSNNKGTRASSYXEPDAXEQWKFPHKKL QLPGXTHE
380	1730	A	3207	187	507	GGTGHPHPARPPLSGVGGCQCSHSKPWTAGS PEQRDHPAPHKQIEAGQGLPGPQAWGG*KGP AXLLPGPGGGPGPVASLEARAQASSGVTPNG GGRTYPYPTFSSGE
381	1731	A	3225	1	840	GTRPGHLPAPSDGFCV/HL*SIPSWGSF*GESL/EMQLITSLGLQEFDIARNVLELIYAQTLVWIGIFFCPLLPFIQMIMLFIMFYSKNISLMMNFQPPSKAWRASQMMTFFIFLLFFPSFTGVLCTLAITIWRLKPSADCGPFRGLPLFIHSIYSWIDTLSTRPGYLWVWIYRNLIGSVHFFFILTLIVLIITYLYWQITEGRKIMIRLLHEQIINEGKDKMFLIEKLIKLQDMEKKANPSSLVLERREVEQQGFLHLGEHDGSLDLRSRRSVOEGNPRA
382	1732	A	3238	256	38	LLMIKVSSTCFSCHLHHHHHHHHHHHHQGHNS LFFSLKSSSNSSTLPVYLSYNIILVFSKCLVFDF LFSNACL
383	1733	A	3241	1542	343	KGAPSFVRLYQYPNFAGPHAALANKSFFKAD KVTMLWNKKATAVLVIASTDVDKTGASYYG EQTLHYIATNGESAVVQLPKNGPIYDVVWNS SSTEFCAVYGFMPAKATIFNLKCDPVFDFGTG PRNAAYYSPHGHILVLAGFGNLILQI*AD/IMK VWNVKNYKLISKPVASDSTYFAWCPDGEHIL TATCAPRLRVNNGYKIWHYTGSILHKYDVPS NAELWQVSWQPFLDGIFPAKTITYQAVPSEVP NEEPKVATAYRPPALRNKPITNSKLHEEEPPQ NMKPQSGNDKPLSKTALKNQRKHEAKKAAK QEARSDKSPDLAPTPAPQSTPRNTVSQSISGDP EIDKKIKNLKKKLKAIEQLKEQAATGKQLEK
384	1734	Ā	3242	3	678	NQLEKIQKETALLQELEDLELGI IRSPAARSPGLETPTCLLFVIAAIAAVFVDSAIP RLTQHRPQDGSFPYTILDPPLYLPGQCAPPQP LSQCARRVHGEKLRRPTFGPRHRGAGTAKMS ASLVRATVRAVSKRKLQPTRAALTLTPSAVN KIKQLLKDKPEHVGVKVGVRTRGCNGLSYTL EYTKTKGDSDEEVIQDGVRVFIEKKAQLTLL GTEMDYVEDKLSSEFVFNNPNIKGTCGCGES FNI
385	1735	A .	3243	3190	664	VAMGTPRAQHPPPPQLLFLILLSCPWIQGLPL KEEEILPEGSETPTVASEALAELLHGALLRR GPEMGYLPGPPLGPEGGEEETTTTIITTTTVTT TVTSPVLCNNNISEGEGYVESPDLGSPVSRTL GLLDCTYSIHVYPGYGIEIQVQTLNLSQEEELL VLAGGGSPGLAPRLLANSSMLGEGQVLRSPT NRLLHFQSPRVPRGGGFRIHYQAYLLSCGFP PRPAHGDVSVTDLHPGGTATFHCDSGYQLQG EETLICLNGTRPSWNGETPSCMASCGGTIHNA TLGRIVSPEPGGAVGPNLTCRWVIEAAEGRRL HLHFERVSLDEDNDRLMVRSGGSPLSPVIYDS DMDDVPERGLISDAQSLYVELLSETPANPLLL SLRFEAFEEDRCFAPFLAHGNVTTTDPEYRPG ALATFSCLPGYALEPPGPPNAIECVDPTEPHW NDTEPACKAMCGGELSEPAGVVLSPDWPQS YSPGQDCVWGVHVQEEKRILLQVEILNVREG DMLTLFDGDGPSARVLAQLRGPQPRRRLLSS GPDLTLQFQAPPGPPNPGLGQGFVLHFKEVPR NDTCPELPPPEWGWRTASHGDLIRGTVLTYQ CEPGYELLGSDILTCQWDLSWSAAPPACQKI

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion MTCADPGEIANGHRTASDAGFPVGSHVQYRC
						LPGYSLEGAAMLTCYSRDTGTPKWSDRVPKC ALKYEPCLNPGVPENGYQTLYKHHYQAGESL RFFCYEGFELIGEVTITCVPGHPSQWTSQPPLC KVTQTTDPSRQLEGGNLALAILLPLGLVIVLG SGVYIYYTKLQGKSLFGFSGSHSYSPITVESDF SNPLYEAGDTREYEVSI
386	1736	A	3250	5725	3984	GTSTVTMATKKHFSIILNLLGMLLKKDNQDT RKLLMTWALEVAVVMKKSETYAPLFCLPSF HKFCKGLLADTLVEDVNICLQACSSLHALSSS LPDDLLQRCVDVCRVQLVHRGTCIRQAFGKL LKSIPLGVFLSNNNHTEIQEISLALRSHMSKAP SNTFHPQDFSD/VISFIL YGNSHRTGKDNWLE RLFYSCQRLDKRDQSTIPRNLLKTDAVLWQW AIWEAAQFTVLSKLRTPLGRAQDTFQTIEGIIR SLAGHTLNPDQDVSQWTTADNDEGHGNNQL RLVLLLQYLENLEKLMYNAYEGCANALTSPP KVIRTFLYTNRQTCQDWLTRIRLSIMRVGLLA GQPAVTVRHGFDLLTEMKTTSLSQGNELEVSI MMVVEALCELHCPEAIQGIAVWSSSIVGKHL LWINSVAQQAEGRFEKASVEYQEHLCAMTG VDCCISSFDKSVLTLASAGCKSASLKHCLNGE SRKSVLSKPTDSSPEVINYLGNKACECYISTA DWAAVQEWQNAIHDLKKSTSSTSLNLKADF NYIKSLSSFESGKFVECTEQLEILPGENINLLA GGSKEKIDMKKLLRNM
387	1737	A	3255	380	76	MDIFLYNCKYQVQTEI*NSIQHIMA\SKKLSRF LKYVHNL*AENYKTLMK*INEDLNKQRDVPY S*TARLNKMSIPTKTIFRFKAIYIKIPATYFIET NMQ
388	1738	A	3260	685	428	PQWLGLQVYALPPANFVFFVEMRSTILAQTG FELLDSSDLPASASKSAGITCMSHHARTLSLK *WPFCLSATQEKFC*PASEGVAW
389	1739	A	3269	1	332 ·	LDGYHTPIYMLNRIIRLPAAL*IISDQTGHALTI LTRLETQMINADYQNKLTLDYLLTTDREVYE PFNLTNYCLHIHNQRLGAYDLG*V*Q/KLAHV PVQV*HGFDPEAMFR
390	1740	A	3270	2	372	GRCHDQNKGKSDGPDAQAEACGGESTYQEL LVNQNPIGQPLACRRLTRKIYEGIKKAVKPNH SPRGVKKVHKFVNKGEKGIMVLAGDTLGIGV YCLLPCMC*DRKLTYAHIPSTTDLGAGAGY
391	1741	A	3273	1	187	FFQEMLDIMKAISDMMGKCTYPVLKEDAPRQ HVETFFQIEELTRSQEGMKLGENFLMFAMPP DDSKESKGK*FFQEMLDIMKAISDMMGKCTY PVLKEDAPRQHVETFFQVGINQKSRGHEVRR KFPDVCHAPR
392	1742	A	3281	901	521	FFFGDGVSPCRQAGV*WHDLDSLQNLPPGFK RFSYLSLPSSW\DYRHVLPRQANFCIF/M*RRG FTMLARMVSIS*PRDLPALASQSAGITGVSHH APPQMDFTFALLCFALKGCLPROKEGGTLNLI
393	1743	A	3283	385	3	RNRSVVPEFVLLGLSAGPQTQTLLFVVIC LLTVMGNLLLLVVINADSCLHTPMYFFLGQL SFLDLCHSSVTAPKLLENLLSEKKTISVEGCM A*VFFVFATGGTESSLLAVMAYDRYVAIRTR G
394	1744	A	3284	575	1054	CTKCKADCDTCFNKNFCTKCKSGFYLHLGKC LDNCPEGLEANNHTMECVSIVHCEVSEWNP WSPCTKKGKTCGFKRGTETRVREIIQHPSAKG NLCPPTNETRKCTVQRKKCQKGERGKKGRE

SEQ ID	T CEO ID	1 N/-4	Lego	Dead: -4-3	Danding 4 - 1	Transaction of the second second
	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
eotide	peptide seq-		in USSN	nucleotide location	location	F=Phenylalanine, G=Glycine, H=Histidine,
seq-	uence	!	09/496	correspondi	corresponding to last amino	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
uence	20,100	1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1			'17	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		1	1	peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
	†			,		RKRKKPNKGESKEAIPDSKSLESSKEIPEQREN
}		i				KQQQ
395	1745	A	3286	1	340	RVLYVPSMGFCILVAHGWQKISTKSVFKKLS
İ		1				WICLSMVILTHSLKTFHRNWDWESEYTLFMS
						ALKVNKNNAKLWNNVGHALENEKNFERAL
L		1			l	KYFLQATHVQPDDIGAHMNVGR
396	1746	A	3293	1	172	GFRAVVMTVKTEAAKGTLTYSRMRGMVAIL
		1			'	IAFMKQRRMGLNDFIQKIANNSYACKQ
397	1747	Α	3295	12	401	AEPACGASSCTPPSLRSSSSQSVGPLRPGRPL
[ĺ				ĺ	WSEACAFL*AAAPQGPASPCCGLPSGFPRVW
1		ļ				AQCCPPGGALRFPEGLGSVLSPRRCPQVSRGS
1		Ì				GLSAVPQEVPSGFLGPGLRACPQEAPSRFLRA
		<u> </u>			}	GLT
398	1748	Α	3300	1912	2768	KQRRWQNIQRKGPKRYIVIAGNSQSHQPMIFS
	1					MLRKLPKVTCRDVLPEIRAICIEEIGCWMQSY
		1				STSFLTDSYLKYIGWTLHDKHREVRVKCVKA
1	1	1	Į į			LKGLYGNRDLTARLELFTGRFKDWMVSMIV
1	l	ĺ	1			DREYSVAVEAVRLLILILKNMEGVLMDVDCE
		}	[SVYPIV*ASN*GLASAVGEFLYWKLFYPECEI
		İ				RTMGGREQRQSPGAQRTFFQLLLSFFVESKSH
1	1	ĺ				SVTQAGVQWQFSAHRDLCLPGSSNSHVSASR
			,			VAGIAGAHRHTWLIYVFFSWRQGFAVLAGL VSNS
399	1749	A	3301	536	2391	LRSYGCKAPSRISHLHK\FLFLLLPSLLMGYSE
3,,,	1777	^	3301	330	2391	SPPPITDSWAPFISLTHHVLSQSQSPLSSNCWI
		ĺ	i			CLSTHTQ*FTALPADLLTWTQSNVSLHISYLAI
						PFLADSFLKPV/L*PGNSAKHLSFKLSSLSMVS
						GRAVALLHLIASGLTSIQTNTASSKPPIWGY\L
1	1					STOTSFISPPPLCLSRTYPNPAHATMVGOVPO
						SLCGLIFTL/RTPCRPSILHPNYKIISTSAWOKV
						LCFSGSPTIHTSLHLTTGSSFLSFHPIPGFPAAN
						SALYVSSLKGPPGKNVTIPSPVTGT*QPPHRGS
						N/RLTVDKDNFFLSPKPNSLHQLPSQ\TPYQAL
						TGAALAGSYPIWENENTLSWLPTFTYNFCLST
				l		PSLFFLCDTN*YLCLPANWSGTCTLVFQAPTI
		l i		ļ	'	NILPPNQTILISVEASISSSPIRNKWALHLITLLT
					İ	GLGITAALGTGIAGITTSITSYQTLFTTLSNTVE
					٠.	DMHTSITSLQRQLDFLVGVILQNWRVLDLLT
				ļ		TEKGGTCIYLQEECCFCVNESGIVHIAVRRLH
					;	DRAAEL*HQVADSWWQGSSLLRWIPWVAPF
[j	ľ	LGPLIFLFLLLMIGPCIFNLVSRFISQRLNCFIQ
[ASMQKHIDNIFHLCHV*YQSLRGNHSEAPEPR
400	1750	Α	3303	2	452	P TINUDUSCOVINGETT A RIPRIDUS ELA TERRITOR
1400	1730	А	3303	2	453	THWRHSSGVPGSTTARRRRRELEIATSDNQE
	,					YYNRLCQEVTNRERNDQKMLADLDDLNRTK
	.				i.:	KYLEERLIELLRDKDALWQKSDALEFQQKLS AEERWLGDTEANHCLDCKREFSWMVRRHHC
]	·					RICGRIFCYYCCNNYVLSKHGGKKERCC
401	1751	A	3304	1	626	MAPQHSSLDDKVPQQASTVCFEFQDILQHSQ
'*'	.,,,	^	3304	•	020	CTEHKDSL WGPGARSQPFGAHNTRLSPDSCP
						EKIVLRALKDSRAGMPEQDKDPGVQENPDD
	}			1	ł	ORRVPOGTGDAPSAFRPLWDNGGLSPFVSRP
			}	1		GPLERDLHAQRSEVTYNQRSQSSWMSSFPKR
		ļ	Ì	1		NAFVSPYSSMGQAQP/GLPKTNPIGESCCWEG
	ļ	Ì	J	ļ		LSLSTQILG*QKPSKYIPSLCKR
402	1752	A	3305	1678	172	MELPSGPGPERLFDSHRLPGDCFLLLVLLLYA
		1				PVGFCLLVLRLFLGIHVFLVSCALPDSVLRRF
			ŀ	l	1	VVRTMCAVLGLVARQEDSGLRDHSVRVLISN
					ļ	HVTPFDHNIVNLLTTCSTVSESEAESATGRFP
				i		

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion GAQLKAPLSPLAFRMEDTEALPLTPILYPTCQ FFFFUFLNIFILLAFSSPGSQPLLNSPPSFVCWSR GFMEMNGRGELVESLKRFCASTRLPPTPLLLF PEEEATNGREGLLRFSSWPFSIQDVVQPLTLQ VQRTLVSVTVSDASWVSELL\WSLFVPFTVY QVRWLRPVHRQLGEANEEFALRVQQ\LVAKE LG\QTGTRLTPA\DKAEHMKRQRHPRLRPQS AQSSFPPSPWVLSS/SDVQTGQTLGFREFKESF CPHVAIGVFIPERPWPKTGCCKTLTIHLILL*G GPVSFSCPEDIHPRGT*VPTQQASGLPSFPSYG PARGGVL*HPSAQQPLTFA\KSS\WARAGRAL
	1	[1		QERKQ\ALYEYARRFTERRAPGGLD
403	1753	Α	3307	44	447	DPSPSLLAVALGLRAGERTRSGPGSSSPSGGIS GGASAGLASSPECACGRSHFTCAVSALGECT CIPAQWQCDGDNDCGDHSDEDGCILPTCSPL DFHCDNGKCIRRSWVCDSDNDCEDDSDEQD CPPRECEED
404	1754	A	3311	409	1	PRHGWGRRVLGRDRPRLQKVKKSVKAIYIPG QDHVQNEEIYARVLDKFGSNFLSRDNADLGT AFVKFSTLTK*LSALLKNLLQGLSRNVIFTLDS LLKGDLKGVKGDLKKPFDKAWKDYETKFAK IEKEKREREWR
405	1755	A	3322	12	458	AAVPVENPWDDPRVRPRVRIFTWEDCIAGQA KVLCNDSYGVTIDWSPKGAFIRLTSQSVGNG HPASKENDQMVDTIKNTTKVPIIWTYGDMVE PRPQMIRPAVGAKHKELWKILMALKKIK\IWE GKYTKPSQYNPNYMLELAHNDSVW
406	1756	A	3324	-:	426	LSMLSTISTEHRLSVLWPIWYCCHCPTHLSAV MCVLLWALSLLQSILEWMFCSFLFSDVDSDN WCQILDFLTAVWLIFLILVLCGFTLVLLVRIIC GSQKMPLTRLYVTILLTGLVFLFCSLPLSIQ*F LLYWIEKDLDDL
407	1757	A	3328	213	1841	SGDLSPAELMMLTIGDVIKQLIEAHEQGKDID LNKVKTKTAAKYGLSAQPRLVDIIAAVPPQY RKVLMPKLKAKPIRTASGIAVVAVMCKPHRC PHISFTGNICVYCPGGPDSDFEYSTQSYTGYEP TSMRAIRARYDPFLQTRHRIEQLKQLGHSVD KVEFIVMGGTFMALPEEYRDYFIRNLHDALS GHTSNNIYEAVKYSERSLTKCIGITIETRPDYC MKRHLSDMLTYGCTRLEIGVQSVYEDVARD TNRGHTVKAVCESFHLAKDSGFKVVAHMMP DLPNVGLERDIEQFTEFFENPAFRPDGLKLYP TLVIRGTGLYELWKSGRYKSYSPSDLVELVA RILALVPPWTRVYRVQRDIPMPLVSSGVEHG NLRELALARMKDLGIQCRDVRTREVGIQEIH HKVRPYQVELVRRDYVANGGWETFLSYEDP DQDILIGLLRLRKCSEETFRFELGGGVSIVREL HVYGSVVPVSSRDPTKFQHQGFGMLLMEEA ERIAREEHGSGKIAVISGVGTRNYYRKIGYRL QGPYMVKMLK
408	1758	A	3335	3	467	AIASPRAAGIRHELTSTMAAGKNKRLTKGGK KGAKKKAV/DNIINIGKTLVTRTQRTKIASDG LKGRVFEESLADLQND\TDGYLLRVI*VAFTT ERTNQI/REVFNKLIPDSIGKDIEKACQSIYPLH DDFARKVKMLKKPKFELRKLMELHGEGSS
409	1759	A	3338	7	1252	PRWRNSARDEILLSFPQNYYIQWLNGSLIHGL WNLASLFSNLCLFVLMPFAFFFLESEGFAGLK KGIRARILETLGMLLLLALLILGIVWVASALID NDAASMESLYDLWEFYLPYLYSCISLMGCLL

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LLLCTPVGL\SRMFTVMGQLLVKPTILEDLDE QIYIITLEEEALQRPTKWAVFIRW/KYNIMELE QELENVKTLKTKLERRKKASAWERNLVYPA VMVLLLIETSISVLLVACNILCLLVDETAMPK GTRGPGIGNASLSTFGFVGAALEIILTYLMVS SVVGFYSLRFFGNFTPKKDDTTMTKIIGNCVS ILVLSSALPVMSRTLGITRFDLLGDFGRFNWL GNFYIVLSYNLLFAIVTTLCLVRKFTSAVREE LFKALGLHKLHLPNTSRDSETAKPSVNGHQK
410	1760	A	3339	127		GSHRFSLASPLDPEVGPYCDTPTMRTLFNLL WLALACSPVHTTLSKSDAKKAASKTLLEKSQ FSDKPVQDRGLVVTDLKAESVVLEHRSYCSA KARDRHFAGDVLGYVTPWNSHGYDVTKVFG SKFTQISPVWLQLKRRGREMFEVTGLHDVDQ GWMRAVRKHAKGLIP*CLGSCLRTGLTMISG/ YVLDSEDEIEELSKTVVQVAKNQHFDGFVVE VWNQLLSQKRVGLIHMLTHLAEALHQARLL ALLVIPPAITPGTDQLGMFTHKEFEQLAPVLD GFSLMTYDYSTAHQPGPNAPLSWVRACVQV LDPKSKWRSKILLGLNFYGMDYATSKDAREP VVGARYIQTLKDHRPRMVWDSQVSEHFFEY KKSRSGRHVVFYPTLKSLQVRLELARELGVG VSIWELGQGLDYFYDLL*VGIAASAVDVFFSK PWSE
411	1761	Α	3342	74	2701	VATRKLAKGFTQFAKMTEGTKKTSKKFKFFK FKGFGSFSNLPRSFTLRRSSASISRQSHLEPDTF EATQDDMVTVPKSPPAYARSSDMYSHMGTM PRPSIKKAQNSQAARQAQEAGPKPNLVPGGV PDPPGLEAAKEVMVKATGPLEDTPAMEPNPS
412	17/0		22.47			AVEVDPIRKPEVPTGDVEEERPPRDVHSERAA GEPEAGSDYVKFSKEKYILDSSPEKLHKELEE ELKLSSTDLRSHAWYHGRIPREVSETLVQRN GDFLIRDSLTSLGDYVLTCRWRNQALHFKIN KVVVKAGESYTHIQYLFEQESFDHVPALVRY HVGSRKAVSEQSGAIIYCPVNRTFPLRYLEAS YGLGQGSSKPASPVSPSGPKGSHMKRRSVTM TDGLTADKVTRSDGCPTSTSLPRPRDSIRSCA LSMDQIPDLHSPMSPISESPSSPAYSTVTRVHA APAAPSATALPASPVARRSSEPQLCPGSAPKT HGESDKGPHTSPSHTLGKASPSPSLSSYSDPDS GHYCQLQPPVRGSREWAATETSSQQARSYGE RLKELSENGAPEGDWGKTFTVPIVEVTSFNP ATFQSLLIPRDNRPLEVGLLRKVKELLAEVDA RTLARHVTKVDCLVARILGVTKEMQTLMGV RWGMELLTLPHGRKLRLDLLERFHTMSIML AVDILGCTGSAEERAALLHKTIQLAAELRGT MGNMFSFAAVMGALDMAQISRLEQTWVTLR QRHTEGAILYEKKLKPFLKSLNEGKEGPPLSN TTFPHVLPLITLECDSAPPEGPEPWGSTEHGV EVVLAHLEAARTVAHHGGLYHTNAEVKLQG FQARPELLEVFSTEFQMRLLWGSQGASSSQA RRYEKFDKVLTALSHKLEPAVRSSEL
412	1762	A	3347	1	898	IDRAAECRTKPLPMAVSIRGNADSIVACLVLM VLYLIKKRLVACAAVFYGFAVHMKIYPETYI LPITLHLLPDRDNDKSLRQFRYTFQACL*ELL KRLCNRTALMFVAVAGLTFFALSFGFYYEYG WEFLEHTYFYHLTRRDIRHNFSPYFYMLYLT AESKWSFSLGIAAFLPQLILLSAVSFAYYRDL VFCWFLHTSIFVTFNKVCTSQYFLWYLCLLPL

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion VMPLVRMPWKRAVVLLMLWFIGQAMWLAP AYVLEFQGKNTFLFIWLAGLFFLLINCSILIQII SHYKEEPLTERIKYD
413	1763	A	3361	3	474	PIPVRWNSLEGRLLRGYEQHANDGKDYISRN *DLRSWTAADMAAQITKRKWEAEEFAEQIKA YLEGTCVER/LRTHLENGKETLQLTEQSSQPTI PIVGIVAGLVLLGAVVTGAVVSAVMCRKKNS GHFLPTDRVSYSEAASSDHAQGSDVSLTACK V
414	1764	A	3363	1488	453	HQILELKKKILKTYNPDYDEDLVQEASSEDVL GVHMVDKDTERDIEMKRQLRRLRELHLYST WKKYQEAMKTSLGVPQRERDEGSLGKPLCP PEILSETLPGSVKKRVCFPSEDHLEEFIAEHLP EASNQSLLTVAHADAGTQTNGDLEDLEEHGP GQTVSEEATEVHMMEGDPDTLAELLIRDVLQ ELSSYNGEEE\DPEEVKTSLGVPQRGDLEDLE EHVPGQTVSEEATGVHMMQVDPATLAKSDL EDLEEHVPEQTVSEEATGVHMMQVDPATLA KQLEDSTITGSHQQMSASPSSAPAEEATEKTK VEEEVKTRKPKKKTRKPSKKSRWNVLKCWD IFNIF
415	1765	Α	3369	431	315	IPWSWVGRLSVRKMSILF*LTYNYNAILNKTP PSFSPSL
416	1766	A	3373	42	651	RQEKMGLGEIGASGVLRSMLKERKKQNMKG NGNVTLTPLLPAVQCGCHLQPAGRSPLPSSHS APGLCSPLHPLQPQQEASTCPSGTLQGREKAA PGQGRPLCSLWAGGAGA\PGERGAEGRGPSD QAPDPKSGPWLFPPGLGAPAEVRLHNVPHNL RRPPLP*ARGK*PPNSGCPWSEGRAKQPLSCG PKPQCSLPSQVPGDTH
417	1767	A	3382	2	2061	EAQDPRACGPDAGGRFAARDAPGNSLRPPPS SPP/GWPGQLRLLPRVPGSELRCGKPERGRLP ASPPGKIRGWPPGISKRPGLGGRSFPPGFAPRT WRPEARGPSVQSLPPIFSPQSAQTTAR*RPGAP KNAGRCGGA\RGPRLSLGPPPGPPPAPALPAR ASAGAGAAAALAVGGVRGAGGARGTGGY GHCSGR/PTGRTGPGPQGPGPPMPARPR*AS\S TRGSRRGPGSRPARAAAAPRAGDHGRRPVRV HLRQHTAV*EPRLGDATAPPGGAAGPGAPAP R\GPGWDCALLPSPGPRSPRAVGCAEPEIWDP SPRRGTSPVPSVRSLRSEPANPRLGLPALLNSY PLKGPGLPPPWGPRTQTGHVIITVQPSGSCIEH SKSLD/RGPWGAPPWGPSSSGLCSPKLATAGP PQSWGLCQIGRRRGLGGPGLKRGET/GLL*GC SMDHANRTKGPGVPTSNRCFSHIPG\GDGCSD HSSCEGHPDLHAGREMPAAPGLSELERVRFT VGCGGLASGISSASVSGLSPNRAGGPGQDW EMYPVSWQTQESGGQG/SPKTGR*VGMLQA GAGSLQGGTGDGVWGLWEDGP/RG*DSPLPS GTGTEP*TPTTSIPFFPQPSGVYPSRATLLPMPS Y*ALGPSANKSEKPLLSFLYRGLCCRISLQLA KGIGQLSEIPLLNVETAFWSMWVTYFRK
418	1768	A	3398	304	2121	EEEEEEDEDDDDNNEEEEFECYPPGMKVQV RYGRGKNQKMYEASIKDSDVEGGEVLYLVH YCGWNVRYDEWIKADKIVRPADKNVPKIKH RKKIKNKLDKEKDKDEKYSPKNCKPPALGPN PPFQTNPISWKWYPKLDLTDAKNSDTAHIKSI EITSILNGLQASESSAEDSEQEDERGAQDMDN NGKEESKIDHLTNNRNDLISKEEQNSSSLLEE

[00 0 10 1						
SEQ ID NO: of	SEQ ID NO: of	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	hod	ID NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid,
eotide	seq-		USSN	location	corresponding	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
			* * * *	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		Ì		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
		١.	ĺ	sequence		nucleotide insertion
						NKVHADLVISKPVSKSPERLRKDIEVLSEDTD
1	ı					YEEDEVTKKRKDVKKDTTDKSSKPQIKRGKR
1 1		Ì				RYCNTEECLKTGSPGKKEEKAKNKESLCMEN
1 [[SSNSSSDEDEEETKAKMTPTKKYNGLEEKRK
1 1					1	SLRTTGFYSGFSEVAEKRIKLLNNSDERLQNS
			l			RAKDRKDVWSSIQGQWPKKTLKELFSDSDTE
1 1						AAASPPHPAPEEGVAEESLQTVAEEESCSPSV
}						ELEKPPPVNVDSKPIEEKTVEVNDRKAEFPSS
						GSNFSA*IPLPYLHLNRLHQSL*QKGSRQQSS
1						VTVSEPLAPNQEEVRSIKSETDSTIEVDSVAGE
						LQDLQSERE*LASRF*CQCELKQ**SARTRTS*
						KSLYRSEKSERCSGRRKFIKKAEKKP*SNSGK
419	1769	A	3399	206	462	QQKEGK
415	1709	A	צעננ	206	463	QRECLSIHIGQAGIQIGDACWELYCLEHGIQP
1 1						NGVVLDTQQDQLENAKMEHTNASFDTFFCE TRAGKHVPRALFVDLEPTVIDGIR
420	1770	A	3408	1010	685	PRI SEEE*INGSVII VEO A DUOVIDIO CORORE
120	1770	Λ	3400	1010	003	RRLSFFF*IWSSVLVTQARVQWRDLGSPQPLP PGFKRFSCLSLPSSWDYRHPSPRPVNF/HVFLV
						VMGFHHVGQAGLELLTSGDLPALASQSARIT
						GVNHCAQPRGHFH
421	1771	A	3409	355	1326	ADSNLIESCWQELGLGPWGGDWRVEQVGAS
			3,0	305	1520	ASLRFPREVCSIRFLFTAVSLLSLFLSAFWLGL
1						LYLVSPLENEPKEMLTLSEYHERVRSQGQQL
						QQLQAELDKLHKEVSTVRAANSERVAKLVF
1						QRLNEDFVRKPDYALSSVGASIDLQKTSHDY
						ADRNTAYFWNRFSFWNYARPPTVILEPHVFP
l i						GNCWAFEGDQGQVVIQLPGRVQLSDITLQHP
						PPSVEHTGGANSAPRDFAVFFLLSFFTHQGLQ
						VYDETEVSLGKFTFDVEKSEIQTFHLQNDPPA
1 1				ł		AFPKVKIQILSNWGHPRFTCLYRVRAHGVRT
400	1772		2410			SEGAEGSAQGPH
422	1//2	A	3412	2	421	EFDAQPSIGALVVFKRP*ATTGSDPGPKRGMN
1 1	1		' f			YLVSCSMRSPESGKGEPGTARDYTPMGRPPP
			Ì	Ì		PVPSVSPGPLPGSLAIAPHSPEPHPWEQQPPRG
i						QARSPPGGWLGSAT/RVRRPHNHP/RGH/HSP
423	1773	A	3420	91	706	VDTAGAPASPGPDVCE
423	1//3	^	3420	91	/06 .	DAQRAIYSSVGPAVSLRQRQQDGAVKESGR/
	ł		1	i		RGGVRSFSRAAAAMAPIKVGDAIPAVEVFEG EPGNKVNLAELFKGKKGVLFGVPGAFTPGCS
					[KTHLPGFVEQAEALKAKGVQVVACLSVNDA
		- 1				FVTGEWGRAHKAEGKVRLLADPTGAFGKET
		ļ	1	İ		DLLLDDSLVSIFGNRRLKRFSMVVQDGIVKA
[ſ	ſ	[ſ	ĺ	LNVEPDGTGLTCSLAPNIISQL
424 .	1774	A	3421	4	7688	RQVTRVGTRVLGSTTAAVFLSVEDDNDNAPQ
		l				FSEKRYVVQVREDVTPGAPVLRVTASDRDKG
[[ſ	į	1	1	ľ	SNAVVHYSIMSGNARGQFYLDAQTGALDVV
	ľ	ĺ				SPLDYETTKEYTLRVRAQDGGRPPLSNVSGL
		l	İ]		VTVQVLDINDNAPIFVSTPFQATVLESVPLGY
				i		LVLHVQAIDADAGDNARLEYRLAGVGHDFP
	ŀ	}	}			FTINNGTGWISVAAELDREEVDFYSFGVEAR
		į				DHGTPALTASASVSVTALDVNDNNPTFTQPE
I I]	YTVRLNEDAAVGTSVVTVSAVDRDAHSVITY
1			4			OUTCONTON TO POST OF CONTROL AND THE TANK OF THE PARTY OF THE
·			1	l	l	QITSGNTRNRFSITSQSGGGLVSLALPLDYKLE
						RQYVLAVTASDGTRQDTAQIVVNVTDANTH
						RQYVLAVTASDGTRQDTAQIVVNVTDANTH RPVFQSSHYTVNVNEDRPAGTTVVLISATDE
						RQYVLAVTASDGTRQDTAQIVVNVTDANTH RPVFQSSHYTVNVNEDRPAGTTVVLISATDE DTGENARITYFMEDSIPQFRIDADTGAVTTQA
						RQYVLAVTASDGTRQDTAQIVVNVTDANTH RPVFQSSHYTVNVNEDRPAGTTVVLISATDE DTGENARITYFMEDSIPQFRIDADTGAVTTQA ELDYEDQVSYTLAITARDNGIPOKSDTTYLEI
						RQYVLAVTASDGTRQDTAQIVVNVTDANTH RPVFQSSHYTVNVNEDRPAGTTVVLISATDE DTGENARITYFMEDSIPQFRIDADTGAVTTQA

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
	l .	1100	I .	,		
nucl- eotide	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
1	seq-	Ì	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1	Į.	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
			1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	J			peptide		/=possible nucleotide deletion, \=possible
	ł		ł	sequence		nucleotide insertion
				Sequence		VESTSGIVRTLRRLDRENVAQYVLRAYAVDK
						CMDD A DEDA CENTERIA DA ALDA DEDA DESCRIPTION DE LA COMPANSION DE LA COMPA
ļ	ļ)			GMPPARTPMEVTVTVLDVNDNPPVFEQDEFD
	i		[VFVEENSPIGLAVARVTATDPDEGTNAQIMY
		i				QIVEGNIPEVFQLDIFSGELTALVDLDYEDRPE
						YVLVIQATSAPLVSRATVHVRLLDRNDNPPV
	1]			LGNFEILFNNYVTNRSSSFPGGAIGRVPAHDP
						DISDSLTYSFERGNELSLVLLNASTGELKLSR
						ALDNNRPLEAIMSVLVSDGVHSVTAQCALRV
1						
1]			TIITDEMLTHSITLRLEDMSPERFLSPLLGLFIQ
						AVAATLATPPDHVVVFNVQRDTDAPGGHILN
						VSLSVGQPPGPGGGPPFLPSEDLQERLYLNRS
1						LLTAISAQRVLPFDDNICLREPCENYMRCVSV
1 .			j J			LRFDSSAPFIASSSVLFRPIHPVGGLRCRCPPGF
1						TGDYCETEVDLCYSRPCGPHGRCRSREGGYT
						CLCRDGYTGEHCEVSARSGRCTPGVCKNGGT
1						CVNLLVGGFKCDCPSGDFEKPYCQVTTRSFP
1						AHSFITFRGLRQRFHFTLALSFATKERDGLLL
						YNGRFNEKHDFVALEVIOEOVOLTFSAGEST
	1					
	l i					TTVSPFVPGGVSDGQWHTVQLKYYNKPLLG
						QTGLPQGPSEQKVAVVTVDGCDTGVALRFGS
						VLGNYSCAA\QGTQGGSKKSLDLTGPLLLGG
1 1		١				VPDLPESFPVRMRQFVGCMRNLQVDSRHIDM
						ADFIANNGTVPGCPAKKNVCDSKTCHNGGTC
						VNQWDAFSCECPLGFGGKSCAQEMANPQHF
						LGSSLVAWHGLSLPISQPWYLSLMFRTRQAD
(i i				. 1	GVLLQAITRGRSTITLQLREGHVMLSVEGTGL
						QASSLRLEPGRANDGDWHHAQLALGAIGGP
						GHAILSFDYGQQRAEGNLGPRLHGLHLSNITV
			Ì			GGIPGPAGGVARGFRGCLQGVRVSDTPEGVN
j i	·					· · ·
						SLDPSHGESINVEQGCSLPDPCDSNPCPANSY
						CSNDWDSYSCSCDPGYYGDNCTNVCDLNPC
	,]			EHQSVCTRKPSAPHGYTCECPPNYLGPYCET
İ			I			RIDQPCPRGWWGHPTCGPCNCDVSKGFDPDC
			I			NKTSGECHCKENHYRPPGSPTCLLCDCYPTG
i l			I		-	SLSRVCDPEDGQCPCKPGVIGRQCDRCDNPF
]]			J			AEVTTNGCEVNYDSCPRAIEAGIWWPRTRFG
						LPAAAPCPKGSFGTAVRHCDEHRGWLPPNLF
	1					NCTSITFSELKGFAERLQRNESGLDSGRSQQL
]		ļ	1	1		ALLLRNATQHTAGYFGSDVKVAYQLATRLL
 		1	ļ			AHESTQRGFGLSATQDVHFTENLLRVGSALL
1 1			İ	į		DTANKRHWELIQQTEGGTAWLLQHYEAYAS
1	İ	l	ľ	1	1	
		ŀ				ALAQNMRHTYLSPFTIVTPNIVISVVRLDKGN
[]	l			ļ		FAGAKLPRYEALRGEQPPDLETTVILPESVFR
1	İ			ŀ		ETPPVVRPAGPGEAQEPEELARRQRRHPELSQ
1 1	Ī	- 1		- 1	1	GEAVASVIIYRTLAGLLPHNYDPDKRSLRVPK
1		l	1		i	RPIINTPVVSISVHDDEELLPRALDKPVTVQFR
		Į		1	i	LLETEERTKPICVFWNHSILVSGTGGWSARGC
			ĺ	İ	l	EVVFRNESHVSCQCNHMTSFAVLMDVSRRE
[ĺ	1		ĺ	1	NGEILPLKTLTYVALGVTLAALLLTFFFLTLL
1		j			İ	RILRSNQHGIRRNLTAALGLAQLVFLLGINQA
1				l	İ	DLPFACTVIAILLHFLYLCTFSWALLEALHLY
]	j	J			}	RALTEVRDVNTGPMRFYYMLGWGVPAFITG
	}	- 1		· 1		
1 1			1			LAVGLDPEGYGNPDFCWLSIYDTLIWSFAGP
] [1	l	1	VAFAVSMSVFLYILAARASCAAQRQGFEKKG
1			j	j	ļ	PVSGLQPSFAVLLLLSATWLLALLSVNSDTLL
1	.	1	1			FHYLFATCNCIQGPFIFLSYVVLSKEVRKALK
1 . 1			!	ł		LACSRKPSPDPALTTKSTLTSSYNCPSPYADG
1 1			-	ł	.	RLYQP\YGDSAGSLHSTSRSGKSQPSYIPFLLR
		ļ			•	EESALNPG\QGPPGLGGIPGR/LCFLGRFKDQQ
						H\DS*TRDFDSDLSLEDDQSGSYASTHSSDSEE

	SEQ ID	SEQ.ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
	NO: of	NO: of	hod	ID NO:	beginning	писleotide	D=Aspartic Acid, E=Glutamic Acid,
	nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
	eotide seq-	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
	uence	uence		09/496 914	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
	delice		1	914	ng to first amino acid	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1			residue of	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan,
					peptide	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
				ļ	sequence		nucleotide insertion
			 	 	sequence		EEEEEEEAAFPGEQGWDSLLGPGAERLPLHS
		}	ł	ļ		ł	TPKDGGPGPGKAPWPGDFGTTAKESSGNGAP
i			Ì				EERLRENGDALSREGSLGPLPGSSAQPHKGIL
			İ				KKKCLPTISEKSSLLRLPLEQCTGSSRGSSASE
į			i				GSRGGPPSRPPPRQSLQEQLNGVMPIAMSIKA
ı			[·	GTVDEDSSGSEFLFFNFLH
ļ	425	1775	A	3429	155	1417	GEPAVQSCDCGCTQRSCPWLLVAPGLLSSSSS
i				'		;	RAASVREAEDAPLQPASIHPVSQGSRGPEGSL
							GSAECLPGDPLGARRATRAHSPVPGPPPSLPA
							AGTAVKRGLQPG*GA/GATSTPGTGAATGGL
			j	1		,	CGPAWAAPSAVGPCCCCPSISTTPSQMRSARP
١							SLGCLPSWAS\PGTEHPPGPQGPGPS*DLCSV*
١							KREFQRGPWAGMVILHRISAADPARAPGPDS
١							NLQSALQQPATGCSEPAAVYSPPIGLWGA**P
١]	ļ		EYG*PQHSLPG*TAPADR*P\AGIKDRVYSNSI
1	ſ						YELLENGQRAGTCVLEYATPLQTLFAMSQYS
1					:	•	QAGFSREDRLEQAKLFCRTLEDILADAPESQN
١							NCRLIAYQEPADDSSFSLSQEVLRHLRQEEKE
١							EVTVGSLKTSAVPSTSTMSQEPELLISGMEKP
ŀ	426	1776		2424			LPLRTDFS
ł	420	1776	A	3431	1662	369	AIWWLSWLQHDLLPTPTQVAIDFTASNGDPR
İ							SSQSLHCLSPRQPNHYLQALRAVGGICQDYD/
1							SVGESGAGGNRQGGLAQRIPQLFLLPSDKRFP
ı							AFGFGARIPPNFEVG*MRGKEGDGGRVSQAE
ľ	1			}			KAGPHCSRLALTG\SHDFAINFDPENPECEGK
		i					RGDFHLPRLPADTLHTGAQTPLPRAQLPVPST
	ŀ						HPRPVFI\EISGVIASYRRCLPQIQLYGPTNVAP IINRVAEPAQREQSTGQATKYSVLLVLTDGV
١]						VSDMAETRTAIVRASRLPMSIIIVGVGNADFS
١	ł						DMRLLDGDDGPLRCPRGVPAARDIVQFVPFR
4							DFKDVSPPGPFRLKDSSASHPPKSDLRLPPFD
1		}					VLLRTREPSWPP*SPTSPSDDPASPTLPLTPNHI
ı					į		TVPTL\AAPSALAKCVLAEVPRQVVEYYASQ
٠L							GISPGAPRPCTLATTPSPSP
1	427	1777	A	3446 .	79	9748	GCQSCWPAWPRLRRRGPASAGARLGRKAPW
	-						GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA
-					ļ		ASRPEASGDCRAGRETAMATLEKLMKAFESL
1	l					.	KSFQQQQQQQQQQQQQQQQQQQPPPP
1	ļ	ļ		ļ			PPPPPPQLPQPPPPQAQPLLPQPQPPPPPPPPPPPPPPP
			j			,	GPAVAEEPLHRPKKELSATKKDRVNHCLTIC
	İ			ľ	}		ENIVAQSVRNSPEFQKLLGIAMELFLLCSDDA
							ESDVRMVADECLNKVIKALMDSNLPRLQLEL
]	ļ	- 1	j	ļ	}	YKEIKKNGAPRSLRAALWRFAELAHLVRPQK
1]	l	1	1	CRPYLVNLLPCLTRTSKRPEESVQETLAAAVP
		1	1				KIMASFGNFANDNEIKVLLKAFIANLKSSSPTI
1		}				.]	RRTAAGSAVSICQHSRRTQYFYSWLLNVLLG
1				j		i	LLVPVEDEHSTLLILGVLLTLRYLVPLLQQQV
	.		j		i	İ	KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEL
	1	ĺ	1		ľ	ŀ	TLHHTQHQDHNVVTGALELLQQLFRTPPPEL
	1]	. !	LQTLTAVGGIGQLTAAKEESGGRSRSGSIVELI
					ŀ		AGGGSSCSPVLSRKQKGKVLLGEEEALEDDS
l	1		j	İ			ESRSDVSSSALTASVKDEISGELAASSGVSTPG
	1		- 1	- 1	ŀ	1	SAGHDIITEQPRSQHTLQADSVDLASCDLTSS
	İ		- 1		ŀ		ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG
	l		- 1			.	TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD
1			l	1	ļ		GTDNQYLGLQIGQPQDEDEEATGILPDEASEA
		ľ	- 1	İ			FRNSSMALQQAHLLKNMSHCRQPSDSSVDKF
	į.						
l			}	J	;		VLRDEATEPGDQENKPCRIKGDIGQSTDDDS
		ŀ			;	ļ	APLVHCVRLLSASFLLTGGKNVLVPDRDVRV SVKALALSCVGAAVALHPESFFSKLYKVPLD

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	Soquenco	/=possible nucleotide deletion, \=possible
1				sequence		nucleotide insertion
		_		Bequence		TTEYPEEQYVSDILNYIDHGDPQVRGATAILC
						GTLICSILSRSRFHVGDWMGTIRTLTGNTFSL
					i	ADCIDI I DETLEDESCUTORI ACTAUDACUA
)						ADCIPLLRKTLKDESSVTCKLACTAVRNCVM
1						SLCSSSYSELGLQLIIDVLTLRNSSYWLVRTEL
1						LETLAEIDFRLVSFLEAKAENLHRGAHHYTGL
						LKLQERVLNNVVIHLLGDEDPRVRHVAAASL
i i						IRLVPKLFYKCDQGQADPVVAVARDQSSVYL
						KLLMHETQPPSHFSVSTITRIYRGYNLLPSITD
1						VTMENNLSRVIAAVSHELITSTTRALTFGCCE
						ALCLLSTAFPVCIWSLGWHCGVPPLSASDESR
						KSCTVGMATMILTLLSSAWFPLDLSAHQDAL
						ILAGNLLAASAPKSLRSSWASEEEANPAATK
1 !		1				QEEVWPALGDRALVPMVEQLFSHLLKVINIC
1 1						AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK
1.						GKEKEPGEQASVPLSPKKGSEASAASRQSDTS
1 1						GPVTTSKSSSLGSFYHLPSYLKLHDVLKATHA
						NYKVTLDLQNSTEKFGGFLRSALDVLSQILEL
!		1				ATLQDIGKCVEEILGYLKSCFSREPMMATVC
1 1						VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA
j		1				QRLGSSSVRPGLYHYCFMAPYTHFTQALADA
1		1				SLRNMVQAEQENDTSGWFDVLQKVSTQLKT
j i		į				NLTSVTKNRADKNAIHNHIRLFEPLVIKALKQ
!						YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL
!						DSDQVFIGFVLKQFEYIEVGQFRESEAIIPNIFF
,						FLVLLSYERYHSKQIIGIPKIIQLCDGIMASGR
			i			KAVTHAIPALQPIVHDLFVLRGTNKADAGKE
1			i			LETQKEVVVSMLLRLIQYHQVLEMFILVLQQ
1 1	-	1	Ì			CHKENEDKWKRLSRQIADIILPMLAKQQMHI
				İ		DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF
1 1	Ì	j				VTPNTMASVSTVQLWISGILAILRVLISQSTED
						IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE
1		i				EHSEGKQIKNLPEETFSRFLLQLVGILLEDIVT
1 ;	1		1	,	,	KQLKVEMSEQQHTFYCQELGTLLMCLIHIFKS
		- 1				GMFRRITAAATRLFRSDGCGGSFYTLDSLNLR
1 1						ARSMITTHPALVLLWCOILLLVNHTDYRWW
		ŀ				AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA
	-					AKLGMCNREIVRRGALILFCDYVCQNLHDSE
]	j		ļ			HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS
[[- 1	. [1	[AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI
		ł				HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL
	-	į				ACRRVEMLLAANLQSSMAQLPMEELNRIQEY
	İ	1	1			LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS
		ŀ	i	1		
	1	}	ı	ł		PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK
	1			ļ	ļ	SQCWTRSDSALLEGAELVNRIPAEDMNAFM
	ŀ	-	ĺ	l	ļ	MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA
	İ	1	ŀ		1	AREVTLARVSGTVQQLPAVHHVFQPELPAEP
	l	}		1		AAYWSKLNDLFGDAALYQSLPTLARALAQY
					i	LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL
				l	İ	SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL
				l		WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG
	ļ					EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI
	1		}	1	ļ	TAACEMVAEMVESLQSVLALGHKRNSGVPA
	- 1	Ì	1	Í	1	FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG
	.	į	į	1		WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR
				i		INTLGWTSRTQFEETWATLLGVLVTQPLVME
			ļ	l		QEESPPEEDTERTQINVLAVQAITSLVLSAMT
				l	. 1	VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK
	J	J	J	J	j	LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD
					i	PVPSLSPATTGALISHEKLLLQINPERELGSMS
					ļ	YKLGQVSIHSVWLGNSITPLREEEWDEEEEEE
						THE CAME A MEDIAGITI PROPERTY AND THE PROPERTY OF THE PROPERTY

ſ	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
1	NO: of	NO: of,	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
ı	nucl- eotide	peptide		in USSN	nucleotide location	location	F=Phenylalanine, G=Glycine, H=Histidine,
-	seq-	seq- uence		09/496	correspondi	corresponding to last amino	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
١	uence	uonac		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1		ļ			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
١		1			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
١					peptide		/=possible nucleotide deletion, \=possible
L					sequence		nucleotide insertion
1							ADAPAPSSPPTSPVNSRKHRAGVDIHSCSQFL
							LELYSRWILPSSSARRTPAILISEVVRSLLVVS DLFTERNQFELMYVTLTELRRVHPSEDEILAO
							YLVPATCKAAAVLGMDKAVAEPVSRLLESTL
							RSSHLPSRVGALHGILYVLECDLLDDTAKOLI
	1						PVISDYLLSNLKGIAHCVNIHSQQHVLVMCAT
1							AFYLIENYPLDVGPEFSASIIQMCGVMLSGSE
							ESTPSIIYHCALRGLERLLLSEQLSRLDAESLV
1							KLSVDRVNVHSPHRAMAALGLMLTCMYTG
1							KEKVSPGRTSDPNPAAPDSESVIVAMERVSVL FDRIRKGFPCEARVVARILPQFLDDFFPPQDIM
Ì							NKVIGEFLSNQQPYPQFMATVVYKVFOTLHS
							TGQSSMVRDWVMLSLSNFTQRAPVAMATWS
1							LSCFFVSASTSPWVAAILPHVISRMGKLEQVD
1						•	VNLFCLVATDFYRHQIEEELDRRAFQSVLEV
-	420	1770		2440	2	430	VAAPGSPYHRLLTCLRNVHKVTTC
	428	1778	A	3449	3	430	NSRPSPSAALVEVLLRSGSTFPHTVSGGWAA WGPWSSCSRDCELGFRVRKRTCTNPEPRNGG
1							LPCVGDAAEYQDCNPOACPVRGAWSCWTS
1							WSPCSASCGGGHYQRTRSCTSPAPSPGEDICL
1							GLHTEEALCATQACPEGWS
ſ	429	1779	A	3464	583	3	DALDRRYLERCHPAAGGWVGEGE*ALCQKT/
1	ł				i		RFSGVLEPPLPSLKDGGRFPAWT*RSCSKSLR
ı							AAFTSQFFPSRRSRASPGSAP\GNGQNLTEQHP
ı							CPGSCDPQVLSASWM*VEHRSKFRPPP*NSTI
1							PPES/RS*QGGTVQTGQHSSGREAGSWRARGR NAGRR*KGGGKIGTKQGAVRARKECRGEMA
1	Ì		- 1			ı	SGETDSE
T	430	1780	A	3473	2802	270	FRMRIFLHCPWNQQMWKIWNLLETSLESCKA
l							HLSIQKLLKER\Q\QLPVFKHRDSIVETLKRHR
+							VVVVAGET\GSGKSTQVPHFLLEDLLLNEWE
١							ASKCNIVCTQPRRISAVSLANRVCDELGCENG PGGRNSLCGYQIRMESRACESTRLLYCTTGV
							LLRKLQEDGLLSNVS/HMFIVDEV\HER\SVQS
							DFLLIILKEILQKRSDLHLILMSATVDSEKFST
					ļ		YFTHCPILRISGRSYPVEVFHLEDIIEETGFVLE
							KDSEYCQKFLEEEEEVTINVTSKAGGIKKYQE
							YIPVQTGAHADLNPFYQKYSSRTQHAILYMN
				. [ļ	PHKINLDLILELLAYLDKSPQFRNIEGAVLIFL PGLAHIQQLYDLLSNDRRFYSERYKVIALHSI
-]		1			LSTQDQAAAFTLPPPGVRKIVLATNIAETGITI
	ļ	ļ					PDVVFVIDTGRTKENKYHESSQMSSLVETFVS
	ļ	}					KASALQRQGRAGRVRDGFCFRMYTRERFEG
ł							FMDYSVPEILRVPLEELCLHIMKCNLGSPEDF
1	Ì	-	ì	ł			LSKALDPPQLQVISNAMNLLRKIGACELNEPK
	1						LTPLGQHLAALPVNVKIGKMLIFGAIFGCLDP VATLAAVMTEKSPFTTPIGRKDEADLAKSAL
	ļ			ļ			AMADSDHLTIYNAYLGWKKARQEGGYRSEI
		1		ĺ			TYCRRNFLNRTSLLTLEDVKQELIKLVKAAGF
	İ		- 1	ľ	1		SSSTTSTSWEGNRASQTLSFQEIALLKAVLVA
		j	1				GLYDNVGKIIYTKSVDVTEKLACIVETAQGK
			l		ļ		AQVHPSSVNRDLQTHGWLLYQEKIRYARVY
	1	1	l		1		LRETTLITPFPVLLFGGDIEVQHRERLLSIDGW
	ĺ	ĺ	1	1			IYFQAPVKIAVIFKQLRVLIDSVLRKKLENPK MSLENDKILQIITELIKTENN
1	431	1781	$\overline{\mathbf{A}}$	3474	1	441	FRPAPGHVQP*GGSSAAAGGGLLSHPRPCQQ
					-		PCPPAPAPSRPRSLGSLGQRVPAALATAAQEL
		j		J			PATLGGDGGKPALTAGEAALPGLHRSGVPAA
	[ſ	1			- [AARC*PCT/SRPT*STLSPTQAAWWCRPSRRQ
			į.	2			QRGEASTGGASGRRCGSCFQV

SEQ ID NO: of NO: of nucleotide eotide sequence wence wence with the period with the period wence wenc	NMQID IK/WKE QTLHPS PGDLGY QQLLHS VWLLL SCSRCS PAASPA GAPPPRP
nucleotide sequence Deptide sequence Decation corresponding uence Decation corresponding uence Decation corresponding uence Decation corresponding uence Decation corresponding uence Decation corresponding to last amino acid residue of peptide sequence Decation corresponding to last amino acid residue of peptide sequence Decation corresponding to last amino acid residue of peptide sequence Decation corresponding to last amino acid residue of peptide sequence Decation corresponding to last amino acid residue of peptide sequence Decation corresponding to last amino acid residue of peptide sequence Decation corresponding to last amino acid residue of peptide sequence Decation corresponding to last amino acid residue of peptide sequence Decation corresponding to last amino acid residue of peptide sequence Decation corresponding to last amino acid residue of peptide sequence Decation corresponding to last amino acid residue of peptide sequence Decation corresponding to last amino acid residue of peptide sequence Decation corresponding to last amino acid residue of peptide sequence Decation corresponding to last amino acid residue of peptide sequence Decation corresponding to last amino acid residue of peptide sequence Decation corresponding to last amino acid residue of peptide sequence Decation corresponding to last amino acid residue of peptide sequence Decation corresponding to last amino acid residue of peptide sequence Decation corresponding to last amino acid residue of peptide sequence Decation corresponding to last amino acid residue of peptide sequence Decation corresponding to last amino acid residue of peptide sequence Decation corresponding to last amino acid residue of peptide sequence Decation corresponding to last amino acid residue of peptide sequence Decation corresponding to last amino acid residue of peptide sequence Decation corresponding to last amino acid residue of peptide sequence Decation corresponding to last amino acid residue	NMQID IK/WKE QTLHPS PGDLGY QQLLHS VWLLL SCSRCS PAASPA GAPPPRP
eotide sequence ue	NMQID IK/WKE QTLHPS PGDLGY QQLLHS VWLLL SCSRCS PAASPA GAPPPRP
sequence uence uence 09/496 914 corresponding to last amino acid residue of peptide sequence 1782 A 3478 416 23 QERLTLPNFKTY/YSS*IIEIAWH**K QWFRRESPEIDLCKYS*LSFDKEAKA CSLFNKWCYKNWM/LHVQKKRI*VQKLK\SKWIKDLNVECRITKLLDQEY SRALNSGSR CLAPCSPQPEKNGMQPLLLLPPLLY SLGAPGESTLLVRTSKLLVGLGLQLL QTRSLLALQLHLTSSAPLAAPTAVCA APRSRCVARPAARTGLPTPAPASPAPAASPVLTASPPLPAASPALAAPTAVLA ASPAPPAASPVLTASPPLPAASPALAA	NMQID IK/WKE QTLHPS PGDLGY QQLLHS VWLLL SCSRCS PAASPA GAPPPRP
uence 914 ng to first amino acid residue of peptide sequence 914 peptide sequence 914 acid residue of peptide sequence 918 peptide sequence 918 peptide sequence 928 peptide sequ	NMQID IK/WKE QTLHPS PGDLGY QQLLHS VWLLL SCSRCS PAASPA GAPPPRP
amino acid residue of peptide sequence A	NMQID IK/WKE QTLHPS PGDLGY QQLLHS VWLLL SCSRCS PAASPA GAPPPRP
residue of peptide sequence 1782 A 3478 416 23 QLRRLTLPNFKTY/YSS*IIEIAWH**K QWFRRESPEIDLCKYS*LSFDKEAKAL CSLFNKWCYKNWM/LHVQKKRI*VQKLK\SKWIKDLNVECRITKLLDQEY SRALNSGSR 1783 A 3504 1876 552 CLAPCSPQPEKNGMQPLLLLPPLLY SLGAPGESTLLVRTSKLLVGLGLQLL QTRSLLALQLHLTSSAPLLAAPTAVC APRSRCVARPAARTGLPTPAPASSPAI PAASPAPAASPVLTASPPLPAASPALAAPASPVLTASPPLPAASPALAAPASPVLTASPPLPAASPALAA	NMQID IK/WKE QTLHPS PGDLGY QQLLHS VWLLL SCSRCS PAASPA GAPPPRP
peptide sequence /=possible nucleotide deletion, \=possible nucleotide insertion 432 1782 A 3478 416 23 QLRRLTLPNFKTY/YSS*IIEIAWH**K QWFRRESPEIDLCKYS*LSFDKEAKAL CSLFNKWC/YKNWM/LHVQKKRI**VQKLK\SKWIKDLNVECRITKLLDQEY SRALNSGSR 433 1783 A 3504 1876 552 CLAPCSPQPEKNGMQPLLLLLPPLLYGENGE SLGAPGESTLLVRTSKLLVGLGLQLL QLAPCSPQPEKNGMQPLLLLLPPLLYGENGE QLAPCSPQPEKNGMQPLLLLPPLLYGENGE QAPRSRCVARPAARTGLPTPAPASSPALAAPTAVC APRSRCVARPAARTGLPTPAPAASPALAAPTAVC APRSRCVARPAARTGLPTPAPAASPALAAPTAVC APRSRCVARPAASPVLTASPPLPAASPALAAPTAVC APRSRCVARPAASPVLTASPPLPAASPALAAPTAVC APRSRCVARPAASPVLTASPPLPAASPALAAPTAVC APRSRCVARPAASPVLTASPPLPAASPALAAPTAVC APRSRCVARPAASPVLTASPPLPAASPALAAPTAVC APRSRCVARPAASPVLTASPPLPAASPALAAPTAVC APRSRCVARPAASPVLTASPPLPAASPALAAPTAVC APRSRCVARPAASPVLTASPPLPAASPALAAPTAVC APRSRCVARPAASPVLTASPPLPAASPALAAPTAVC APRSRCVARPAASPVLTASPPLPAASPALAAPTAVC APRSRCVARPAASPVLTASPPLPAASPALAAPTAVC APRSRCVARPAASPVLTASPPLPAASPALAAPTAVC APRSRCVARPAASPVLTASPPLPAASPALAAPTAVC APRSRCVARPAASPVLTASPPLPAASPALAAPTAVC APRSRCVARPAASPVLTASPPLPAASPALAAPTAVC APRSRCVARPAASPVLTASPPLPAASPALAAPTAVC APRSRCVARPAASPVLTASPPLPAASPALAAPTAVC APRSRCVARPAASPVLTASPPLPAASPALAAPTAVC APRSRCVARPAASPVLTASPPLPAASPALAAPTAVC APRSRCVARPATATATATATATATATATATATATATATATATATATA	NMQID IK/WKE QTLHPS PGDLGY QQLLHS VWLLL SCSRCS PAASPA GAPPPRP
sequence nucleotide insertion 432 1782 A 3478 416 23 QLRRLTLPNFKTY/YSS*IIEIAWH**K QWFRRESPEIDLCKYS*LSFDKEAKAL CSLFNKWCYKNWM/LHVQKKRI*VQ QKLK\SKWIKDLNVECRITKLLDQEY SRALNSGSR 433 1783 A 3504 1876 552 CLAPCSPQPEKNGMQPLLLLPPLLYQUE SLGAPGESTLLVRTSKLLVGLGLQLL QTRSLLALQLHLTSSAPLLAAPTAVC APRSRCVARPAARTGLPTPAPASSPAI PAASPAPAESTAIPQPLILLPKYPPAPQ GAPPPRPAASPSPAASPAPAASPVLTASPPLPA ASPAPPAASPVLTASPPLPAASPALAA	NMQID IK/WKE QTLHPS PGDLGY QQLLHS VWLLL SCSRCS PAASPA GAPPPRP
432 1782 A 3478 416 23 QLRRLTLPNFKTY/YSS*IIEIAWH**K QWFRRESPEIDLCKYS*LSFDKEAKAL CSLFNKWC/YKNWM/LHVQKKRI*VQ QKLK\SK WIKDLNVECRITKLLDQEY SRALNSGSR 433 1783 A 3504 1876 552 CLAPCSPQPEKNGMQPLLLLLPPLLYY SLGAPGESTLLVRTSKLLVGLGLQLL QTRSLLALQLHLTSSAPLLAAPTAVC APRSRCVARPAARTGLPTPAPASSPAI PAASPAPAASTAIPQPLILLFK/PPAPA GAPPPRAASPAPAASPVLTASPPLPA ASPAPAASPVLTASPPLPAASPALAA	IK/WKE QTLHPS PGDLGY QQLLHS VWLLL SCSRCS PAASPA GAPPPRP
QWFRRESPEIDLCKYS*LSFDKEAKAL CSLFNKWC/YKNWM/LHVQKKRI*VC QKLK\SKWIKDLNVECRITKLLDQEY SRALNSGSR 433 1783 A 3504 1876 552 CLAPCSPQPEKNGMQPLLLLLPPLLY SLGAPGESTLLVRTSKLLVGLGLQLL QTRSLLALQLHLTSSAPLLAAPTAVC APRSRCVARPAARTGLPTPAPASSPAI PAASPAPAESTA\PQPLILLPK\PPAPC GAPPPRPAASPSPAASPAPPAASPVLT AASPSPAASPAPPAASPVLTASPPLPA ASPAPPAASPVLTASPPLPAASPALAA	IK/WKE QTLHPS PGDLGY QQLLHS VWLLL SCSRCS PAASPA GAPPPRP
CSLFNKWC/YKNWM/LHVQKKRI*VC QKLK\SKWIKDLNVECRITKLLDQEY SRALNSGSR 433 1783 A 3504 1876 552 CLAPCSPQPEKNGMQPLLLLPPLLYI SLGAPGESTLLVRTSKLLVGLGLQLL QTRSLLALQLHLTSSAPLLAAPTAVC APRSRCVARPAARTGLPTPAPASSPAI PAASPAPAESTANPQPLILLPKP/PPAPC GAPPPRPAASPAPAASPALTASPPLPA ASPAPAASPVLTASPPLPAASPALAA	QTLHPS PGDLGY QQLLHS VWLLL SCSRCS PAASPA GAPPPRP
QKLK\SKWIKDLNVECRITKLLDQEY SRALNSGSR 433 1783 A 3504 1876 552 CLAPCSPQPEKNGMQPLLLLPPLLY SLGAPGESTLLVRTSKLLVGLGLQLL QTRSLLALQLHLTSSAPLLAAPTAVC APRSRCVARPAARTGLPTPAPASSPAP PAASPAPAESTA\PQPLILLPKP\PPAPC GAPPPRPA\SPSPAASPAPPAASPVLT AASPSPAASPAPPAASPVLTASPPLPA ASPAPPAASPVLTASPPLPAASPALAA	PGDLGY QQLLHS VWLLL SCSRCS PAASPA GAPPPRP
SRALNSGSR 433 1783 A 3504 1876 552 CLAPCSPQPEKNGMQPLLLLPPLLYG SLGAPGESTLLVRTSKLLVGLGLQLL QTRSLLALQLHLTSSAPLLAAPTAVC APRSRCVARPAARTGLPTPAPASSPAJ PAASPAPAESTA\PQPLILLPKP/PPAPC GAPPPRPAASPSAASPAPAASPVLT AASPSPAASPAPAASPVLTASPPLPA ASPAPAASPVLTASPPLPAASPALAA	QQLLHS VWLLL SCSRCS PAASPA GAPPPRP
SLGAPGESTLLVRTSKLLVGLGLQLL QTRSLLALQLHLTSSAPLLAAPTAVC APRSRCVARPAARTGLPTPAPASSPAI PAASPAPAESTA\PQPLILLPKP/PPAPC GAPPPRPAASPSAASPAPPAASPVLT AASPSPAASPAPPAASPVLTASPPLPA ASPAPPAASPVLTASPPLPAASPALAA	VWLLL SCSRCS PAASPA GAPPPRP
SLGAPGESTLLVRTSKLLVGLGLQLL QTRSLLALQLHLTSSAPLLAAPTAVC APRSRCVARPAARTGLPTPAPASSPAJ PAASPAPAESTA\PQPLILLPKP/PPAPC GAPPPRPAASPSAASPAPPAASPVLT AASPSPAASPAPPAASPVLTASPPLPA ASPAPPAASPVLTASPPLPAASPALAA	VWLLL SCSRCS PAASPA GAPPPRP
QTRSLLALQLHLTSSAPLLAAPTAVC APRSRCVARPAARTGLPTPAPASSPAJ PAASPAPAESTA\PQPLILLPKP/PPAPC GAPPPRPAASPSPAASPAPPAASPVLT AASPSPAASPAPPAASPVLTASPPLPA ASPAPPAASPVLTASPPLPAASPALAA	SCSRCS PAASPA GAPPPRP
APRSRCVARPAARTGLPTPAPASSPAI PAASPAPAESTA\PQPLILLPKP/PPAPC GAPPPRPAASPSPAASPAPPAASPVLT AASPSPAASPAPPAASPVLTASPPLPA ASPAPPAASPVLTASPPLPAASPALAA	PAASPA GAPPPRP
PAASPAPAESTA\PQPLILLPKP/PPAPC GAPPPRPAASPSPAASPAPPAASPVLT AASPSPAASPAPPAASPVLTASPPLPA ASPAPPAASPVLTASPPLPAASPALAA	SAPPPRP
GAPPPRPAASPSPAASPAPPAASPVLT AASPSPAASPAPPAASPVLTASPPLPA ASPAPPAASPVLTASPPLPAASPALAA	
AASPSPAASPAPPAASPVLTASPPLPA ASPAPPAASPVLTASPPLPAASPALAA	
ASPAPPAASPVLTASPPLPAASPALAA	
ASPPVHVASPPVHTASPPVHVASPPVI	-
VHVASPPVHTASPHVHVASPPVHVAS	
ASPPVHTASPPVHVASPPVHTASPHV	
VHTASPPVHVASPPVHVASPPVHVAS	
ASPPVHVASPPVSCSGDSTS	
QPGAVFPHSLAPSLGGWSHLVAALP	
434 1784 A 3516 142 590 GGVNRPRSETEQVKTPVLISSWDYRH	PPPRPA
SFFVFLV*TGF\TALARMVLISWPCDL	
SAGITGVRHHA\RLLYFEQESHSVTQA	
WHNLGSLQPLSLEDRLSPGVLGCSAL	
RTKFGINMVTSRERGTTRLPKEG	
435 1785 A 3529 1 3161 MSLVRAALEALDELDLFGVKGGPQS'	VIHVLA
DEVQHCQSILNSLLPRASTSKEVDASI	
FPAFAVEDSQLVELTKQEITTKLQGRY	GCCRF
LRDGYKTPKEDPNRL YY/ENPAELKL	
EWPLFWTYFILDGVFSGNAEQVQEYF	
VLIKGKNGVPLLPELYSVPPDRVDEE	YONPHT
VDRVPMGKLPHMWGQSLYILGSLMA	
PGEIDPLNRRFSTVPKPDVVVQVYPSI	
SKSPSHQCTIISIRTTRKITAPVSILAET	
KDKGIYVETIAEVYPIRVQPARILSHIY	SSLEIF
LPFLNSVSGCNNRMKLSGRPYRHMG	VLGTSK
LYDIRKTIFTFTPQFIDQQQFYLALDNI	
MLRTDLSYLCSRWRMTGQPTITFPISH	ISMLDE
DGTSLNSSILAALRKMQDGYFGGARV	/QTGKL
SEFLTTSCCTHLSFMDPGPEGKLYSED	
YDYLESGNWMNDYDSTSHARCGDEN	
DHLLAHTAPHPKLAPTSQKGGLDRFQ	AAVQT
TCDLMSLVTKAKELHVQNVHMYLPT	KLFQA
SRPSFNLLDSPHPRQENQVPSVRVEIH	LPRDO
SGEVDFKALVLQLKETSSLQEQADILY	/MLYT
MKGPDWNTELYNERSATVRELLTELY	YGKVG
EIRHWGLIRYISGILRKKVEALDEACT	
QKHLTVGLPPEPREKTISAPLPYEALT	
SEGDMSISILTQEIMVYLAMYMRTQPO	
MFRLRIGLIIQVMATELAHSLRCSAEE	
MNLSPSAMKNLLHHILSGKEFGVERS	
SNVSPAISIHEIGAVGATKTERTGIMQI	
QSPGTSMTPSSGSFPSAYDQQSSKDSR	
QRRRRLDGALNRVPVGFYQKVWKVL	`
GLSVEGFVLPSSTTREMTPGEIKFSVH	
NRVPQPEYRQLLVEAIL\VLTMLADIE	
IIAVEKIVHIANDLFLQEQKTLGADDT	
PASGICTLLYDSAPSGRFGTMTYLSKA	
VQEFLPHSICAMQ	
436 1786 A 3546 73 393 CP*LTWELLEVKKAEVLQDSLDGRYS	TPSSCL
EQPDSCRPYGRSFYALEEKHVIFSLDV	GETON

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide deletion, \=possible nucleotide insertion KGKGKTIRGI*TFKGRKGGTYQREHDANPLA PXSARSCWMRKG
437	1787		3554	5157	2939	AVRAEPGLEELSSGLRAHSPSATTVCEPEAQG SASGCRYAAHPHWGLGGAAAAGGSWEPQPP RPVCEPAGRKPHPPAAPRSPLLPGSRRRPHA AQPGARARTSPPPASARNMAARPAATLAWSL LLLSSALLREGCRARFVAERDSEDDGEEPVVF PESPLQSPTVLVAVLARNAAHTLPHFLGCLER LDYPKSRMAIWAATDHNVDNTTEIFREWLK NVQRLYHYVEWRPMDEPESYPDEIGPKHWP TSRFAHVMKLRQAALRTAREKWSDYILFIDV DNFLTNPQTLNLLIAENKTIVAPMLESRGLYS NFWCGITPKGFYKRTPDYVQIREWKRTGCFP VPMVHSTFLIDLRKEASDKLTFYPPHQDYTW TFDDIIVFAFSSRQAGIQMYLCNREHYGYLPIP LKPHQTLQEDIENLIHVQIEAMIDKPPMEPSQ YVSVVPKYPDKMGFDEIFMINLKRRKGQGGD RWLRTLYEQEIEVKIVEAVDGKALNTSQLKA LNIEMLPGYRDPYSSRPLTRGEIGCFLSHYSV WKEVIDRELEKTLVIEDDVRFEHQFKKKLMK LMDNIDQAQLDWELIYIGRKRMQVKEPEKA
				ļ		VPNVANLVEADYSYWTLGYVISLEGAQKLV GANPFGKMLPVDEFLPVMYNKHPVAEYKEY YESRDLKAFSAEPLLIYPTHYTGQPGYLSDTE TSTIWDNETVATDWDRTHAWKSRKQSRIYSN AKNTEALPPPTSLDTVPSRDEL
438	1788	A	3563	130	527	IFFNSSSLFCRVFCLFLRWSFTLVAQARVQ*C NLSSLQPLPPGFK*FSCLSPPRS*DYRRPPPRPA NFLYF**RQGFTVLGQAGLELLT/S/GDPPTSA SQSAGITGVSHRAWPVHAISTHISLVKTRPSLT TLG
439	1789	A	3565	446	1834	LLQPAMRKSPGLSDCLWAWILLSTLTGRSY GQPSLQDELKDNTTVFTRILDRLLDGYDNRL RPGLGERVTEVKTDIFVTSFGPVSDHDMEYTI DVFFRQSWKDERLKFKGPMTVLRLNNLMAS KIWTPDTFFHNGKKSVAHNMTMPNKLLRITE DGTLLYTMRLTVR\AECPMAFGRDFPM\D\AH ACPLKFGSYAYTRAEVVYEWTREPARSVVV AEDGSRLNQYDLLGQTVDSGIVQSSTGEYVV MTTHFHLKRKIGYFVIQTYLPCIMTVILSQVSF WLNRESVPARTVFGVTTVLTMTTLSISARNSL PKVAYATAMDWFIAVCYAFVFSALIEFATVN YFTKRGYAWDGKSVVPEKPKKVKDPLIKKN NTYAPTATSYTPNLARGDPGLATIAKSATIEP KEVKPETKPPEPKKTFNSVSKIDRLSRIAFPLL FGIFNLVYWATYLNREPQLKAPTPHQ
440	1790	A	3568	1	350	STSSCFPAAAAAIMREIVHLQAGQCGNQIGAK FWEVISDEHGIDPTGTYHGDSDLQLERINVYY NEATGEAPVPSPTALRGPRGPCLG*RPPVPAG GKYVPRAVLVDMEPGTMDSV
441	1791	A	3569	2	1751	FVAVAGAVSGEPLVHWCTQQLRKTFGLDVS EEIIQYVLSIESAEEIREYVTDLLQGNEGKKGQ FIEELITKWQKNDQELISDPLQQCFKKDEILDG QKSGDHLKRGRKKGRNRQEVPAFTEPDTTAE VKTPFDLAKAQENSNSVKKKTKFVNLYTREG QDRLAVLLPGRHPCDCLGQKHKLINNCLICG RIVCEQEGSGPCLFCGTLVCTHEEQDILRGDS NKSQKLLKKLMSGVENSGKVDISTKDLLPH QELRIKSGLEKAIKHKDKLLEFDRTSIRRTQVI

NO: of nucleotide cotide sequence NO: of nucleotide sequence NO: of nucleotide sequence NO: of nucleotide sequence NO: of nucleotide sequence NO: of nucleotide sequence NO: of nucleotide sequence NO: of nucleotide sequence NO: of nucleotide sequence NO: of nucleotide sequence NO: of nucleotide sequence NO: of nucleotide sequence NO: of nucleotide sequence NO: of nucleotide sequence NO: of peptide sequence NE-Phenylalanine, G=Glycine, H=Hilled sequence N=Phenylalanine, G=Glycine, H=Hilled sequence N=Phenylalanine, G=Glycine, H=Hilled sequence N=Phenylalanine, G=Glycine, H=Hilled sequence N=Phenylalanine, G=Glycine, H=Hilled sequence N=Phenylalanine, G=Glycine, H=Hilled sequence N=Phenylalanine, G=Glycine, H=Hilled sequence N=Phenylalanine, G=Glycine, H=Hilled sequence N=Phenylalanine, G=Glycine, H=Hilled sequence N=Phenylalanine, G=Glycine, H=Hilled sequence N=Phenylalanine, G=Glycine, H=Hilled sequence N=Phenylalanine, G=Glycine, H=Hilled sequence N=Phenylalanine, G=Glycine, H=Hilled sequence N=Phenylalanine, G=Glycine, H=Hilled sequence N=Phenylalanine, S=Sequence N=Methionine, N=Asparagine, P=FQ-Glutamic Acid residue of peptide N=Aspartic Acid, E=Glutamic Acid residue of peptide N=Aspartic Acid, E=Glutamic Acid residue of peptide N=Aspartic Acid, E=Glutamic Acid residue of peptide N=Aspartic Acid, E=Glutamic Acid residue of peptide N=Aspartic Acid, E=Glutamic Acid residue of peptide N=Aspartic Acid, E=Glutamic Acid residue of peptide N=Aspartic Acid, E=Sequence N=Aspartic Acid residue of peptide N=Asparti	istidine, e, Proline, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10
sequence Sequence 09/496	Proline, ne, phan, codon, sssible FLQKREEELR EEENSLAEYH DRSSEEPLGVL
uence 914 ng to first amino acid residue of peptide sequence 914 ng to first amino acid residue of peptide sequence 914 ng to first amino acid residue of peptide sequence 92-Glutamine, R=Arginine, S=Serin T=Threonine, V=Valine, W=Trypto Y=Tyrosine, X=Unknown, *=Stop of /=possible nucleotide deletion, \=po nucleotide insertion DDESDYFASDSNQWLSKLERET ELRHASRLSKKVTIDFAGRKILE SRLDETIQAIANGTLNQPLTKLD VNPNMYQSPPQWVDHTGAASC GLEFNSFQHQLRIQDQEFQEGFI QPWASLLVRGIKRVEGRSWYTI TAKKPSPQEVSELQATYRLLRG PSGCLLGCVDLIDCLSQKQFKEO PFVFICKNPQEMVVKFPIKGNPK	ne, ophan, codon, ssible rLQKREEELR EEENSLAEYH DRSSEEPLGVL
amino acid residue of peptide sequence y=Tyrosine, X=Unknown, *=Stope /=possible nucleotide deletion, \=po nucleotide insertion DDESDYFASDSNQWLSKLERET ELRHASRLSKKVTIDFAGRKILE SRLDETIQAIANGTLNQPLTKLD VNPNMYQSPPQWVDHTGAASG GLEFNSFQHQLRIQDQEFQEGFI QPWASLLVRGIKRVEGRSWYTI TAKKPSPQEVSELQATYRLLRG PSGCLLGCVDLIDCLSQKQFKEG PFVFICKNPQEMVVKFPIKGNPK	ophan, codon, essible TLQKREEELR CEENSLAEYH DRSSEEPLGVL
residue of peptide sequence Y=Tyrosine, X=Unknown, *=Stop of peptide sequence Possible nucleotide deletion, \=possible nucleotide insertion DDESDYFASDSNQWLSKLERET ELRHASRLSKKVTIDFAGRKILE SRLDETIQAIANGTLNQPLTKLD VNPNMYQSPPQWVDHTGAASQ GLEFNSFQHQLRIQDQEFQEGFI QPWASLLVRGIKRVEGRSWYTI TAKKPSPQEVSELQATYRLLRG PSGCLLGCVDLIDCLSQKQFKEQ PFVFICKNPQEMVVKFPIKGNPK	codon, ssible FLQKREEELR EEENSLAEYH DRSSEEPLGVL
sequence nucleotide insertion DDESDYFASDSNQWLSKLERET ELRHASRLSKKVTIDFAGRKILE SRLDETIQAIANGTLNQPLTKLD VNPNMYQSPPQWVDHTGASQ GLEFNSFQHQLRIQDQEFQEGFI QPWASLLVRGIKRVEGRSWYTI TAKKPSPQEVSELQATYRLLRG PSGCLLGCVDLIDCLSQKQFKEG PFVFICKNPQEMVVKFPIKGNPK	TLQKREEELR EEENSLAEYH DRSSEEPLGVL
DDESDYFASDSNQWLSKLERET ELRHASRLSKKVTIDFAGRKILE SRLDETIQAIANGTLNQPLTKLD VNPNMYQSPPQWVDHTGAASQ GLEFNSFQHQLRIQDQEFQEGFI QPWASLLVRGIKRVEGRSWYTI TAKKPSPQEVSELQATYRLLRG PSGCLLGCVDLIDCLSQKQFKEQ PFVFICKNPQEMVVKFPIKGNPK	EEENSLAEYH DRSSEEPLGVL
ELRHASRLSKKVTIDFAGRKILE SRLDETIQAIANGTLNQPLTKLD VNPNMYQSPPQWVDHTGAASQ GLEFNSFQHQLRIQDQEFQEGFI QPWASLLVRGIKRVEGRSWYTI TAKKPSPQEVSELQATYRLLRG PSGCLLGCVDLIDCLSQKQFKEG PFVFICKNPQEMVVKFPIKGNPK	EEENSLAEYH DRSSEEPLGVL
SRLDETIQAIANGTLNQPLTKLD VNPNMYQSPPQWVDHTGAASQ GLEFNSFQHQLRIQDQEFQEGFI QPWASLLVRGIKRVEGRSWYTI TAKKPSPQEVSELQATYRLLRG PSGCLLGCVDLIDCLSQKQFKEG PFVFICKNPQEMVVKFPIKGNPK	DRSSEEPLGVL
VNPNMYQSPPQWVDHTGAASQ GLEFNSFQHQLRIQDQEFQEGFI QPWASLLVRGIKRVEGRSWYTI TAKKPSPQEVSELQATYRLLRG PSGCLLGCVDLIDCLSQKQFKEQ PFVFICKNPQEMVVKFPIKGNPK	
QPWASLLVRGIKRVEGRSWYTI TAKKPSPQEVSELQATYRLLRG PSGCLLGCVDLIDCLSQKQFKEG PFVFICKNPQEMVVKFPIKGNPK	JKKAFRSSGF
TAKKPSPQEVSELQATYRLLRG PSGCLLGCVDLIDCLSQKQFKEG PFVFICKNPQEMVVKFPIKGNPK	
PSGCLLGCVDLIDCLSQKQFKEG PFVFICKNPQEMVVKFPIKGNPK	
PFVFICKNPQEMVVKFPIKGNPK	
	, ,
QGAKKGLMKQNKAV	
442 1792 A 3576 1 2019 MPRSHTGERLCEGKEGSQCAEN	
KTAGVKPYECTICGKAFMRLSS	- 1
AIRAI\EKPYKCKEC\GRAFSLSQ	
TGEKPYKCKQCGKTFIYHQPFQ PYECKQCGKALSCSSSLRVHER	
KQCGKAFSCSSSIRVHERTHTGI	
GKAFIS\TTSVLTHMITHNGDRP	YKCKECGKA
FIFPSFLRVHERIHTGEKPYKCK	QCGKAFRWS
TSIQIHERIHTGEKPYKCKECGK	
HVRVHTGEKPYKCKECGKAFSI HTGEKPYECKKCGKTFNYPLDI	_
EKPYECKECAKTFISLENFRRHN	
KCRDCGKVFIFPSALRTHERTH	
CGKAFSCSSYIRIHKRTHTGEK	
AFIYPTSFQGHMRMHTGEKPYK LHSSFR\RHTRIHNYEKPLEC*Q\	CCKECGKAFS
LKKPMRNAQSDRKLY/KCEK*E	
QSCENSH*REKSCQCK*YRKRD	
PHNHVSVSNGPYR/CGSPIRLYN	
VAVVTP*CSTLFKCLWCWCKR	
DSGRGRWLTPVIPALWEÄKAGG ANTVKPHLY	GSKGQEIKTIL
443 1793 A 3578 287 114 DFYERKFEQFIEGHKQIVNKWR	DLLCSWKRK
LSIIKKSVLQNNL*FSAASMRFQ	
444 1794 A 3582 3335 1909 HLFFSLFLAAMAMTGSTPCSSM	SNHTKERVT
MTKVTLENFYSNLIAQHEEREM	`
MEEEGLKDEEKRLRRSAHARKI	1
RLGLEDFESLKVIGRGAFGEVRI VYAMKILRKADMLEKEQVGHI	•
DSLWVVKMFYSFQDKLNLYLIN	
MTLLMKKDTLTEEETQFYIAET	VLAIDSIHQL
GFIHRDIKPDNLLLDSKGHVKLS	
KAHRTEFYRNLNHSLPSDFTFQI TWKRNRRQLAFSTVGTPDYIAP	
KLCDWWSLGVIMYEMLIGYPPI	
KKVMNWKETLTFPPEVPISEKA	
WEHRIGAPGVEEIKSNSFFEGVI	
AISIEIKSIDDTSNFDEFPESDILK	li i
TDYKNKDWVFINYTYKRFEGLT	I AKGAIPSYM
KAAK	AHOYII FFD
GGSRGKGEHFPYEQEIKFFAKV	
NHRLYFLSAASRPLCSGGHASN	
FCKLGVLVRHRISLFGNDATSIV	` '
LDARTVMKTGLESVKSALRAFI	
KTMENLKQGQFTHTRNQPKGV	
LTSLYALGTSKSIYVERORSALO	
FPVAFLETHLDKHNIYSIYNTKS	
TNVEDVCPŅIPSLEKLMEEIVEL	AESGIRYTQ

LILDEFTITLADILYAFYPLIRRGDYNRAKENER REPREAELIFRMAVEVFIYWSSINFKERE ONFVVONENNMSFLITDITSSKMSKAAVSDE ERKKMKRKORYSMOTSILVAKARLI PIGI ERKKMKRKORYSMOTSILVAKARLI PIGI ERKKMKRKORYSMOTSILVAKARLI PIGI NICAPGOGELIALAKNRSIKDTEDEVADINTIDIT DEPEKTVERVIDIANVI, PINDIT DEPEKTVERVIDIANVI, PINDIT DEPEKTVERVIDIANVI, PINDIT DEPEKTVERVIDIANVI, PINDIT DEPEKTVERVIDIANVI, PINDIT DEPEKTVERVIDIANVI, PINDIT DEPEKTVERVIDIANVI, PINDIT DEPEKTVERVIDIANVI, PINDIT DEPEKTVERVI, PINDIT DEPEKTVERVIDIANVI, PINDIT DEPEKTVERVI, PINDIT DEPEKT		000	1 65 6		 -			
muche coided seg- graph of the contraction of the coided segment of the coided segment of peptide residue of peptide requence per per per per per per per per per pe	-		, -	,		,		
oriesponding uence where the services of peptide provided in services of peptide per services of pepti	- 1		I .	hod	1 .	, , ,	5	
sequence Sequence 9,944 0,15	- 1			Ì			1	
uence 914 ng to first amino acid residue of peptide of peptide per peptide of peptide sequence February			, .		1		,	
amino acid residue of peptide sequence p	ļ	-	Lacince	1			i	1 3
residue of peptide sequence Y=Tyrosine, X=Unknown, =>Stop codon, -possible mucleotide dieton, -yososible mucleotide dieton, -yososible mucleotide dieton, -yososible mucleotide dieton, -yososible mucleotide dieton, -yososible mucleotide dieton, -yososible mucleotide dieton, -yososible mucleotide dieton, -yososible mucleotide dieton, -yososible mucleotide dieton, -yososible mucleotide dieton, -yososible mucleotide dieton, -yososible mucleotide dieton, -yososible mucleotide dieton, -yososible mucleotide dieton, -yososible di	Ì	uchco		ĺ	7.17			
pepide sequence Apposible nucleotide deletion, \timesposible nucleotide insertion	١			İ	i			
Sequence nucleotide insertion	- [}	ļ)		Joquenee	
MPHYMEVILPMICSYMSRWWEHOPENNIEL AEMCCTALNSEIMNTLIGHLIKITVNIGIDI GAWMKRLAVFSQPINKVRPQLLKTEFI.PLM EKLKKAATVVSGDPINKVRPQLLKTEFI.PLM EKLKKKAATVVSGDPINKVRPQLLKTEFI.PLM EKLKKKAATVVSGDPINKVRPQLLKTEFI.PLM EKLKKKAATVVSGDPINKVRPQLKTEFI.PLM EKLKKKAATVSGDPINKVRPQLKTEFI.PLM EKLKKKAATVSGDPINKVRPQLKTEFI.PLM KEPPPE AEEL.FRANVEFILESKMSKAANSD ERKKMRKGDRYSMOTSLIVAALKALLPIGI DIEVENTOMEN DE EKKMRKEGRYSMOTSLIVAALKALLPIGI NICAPDOGELALAKNRFSLKDTEDEVDDIRK NIHLQCKLEPPLAKWOMALYKLLESKKRKAVVACI CLVSEHORSKKAVVHKLLESKGKRKAVVACI RMAPLYNLPRIRAVHLEJCSYEKSWETEEF YFEBLEDLAKGAEPPEDGTTRAVDPI-H LILLPSKTALTEKCKLEEDPLYMAYADIMAK CHDEDDDDGEVKSSPERKEMERCKLLYQQ ARLHDROAAEMVLFILESCKSKENVGRIFTEF YFEBLEDLAKGAEPPEDGTTRAVDPI-H LILLPSKTALTEKCKLEEDPLYMAYADIMAK CHDEDDDDGEVKSSPERKEMERCKLLYQQ ARLHDROAAEMVLFILESCKSKERKEKCKLLYQQ ARLHDROAAEMVLFILESCKSKERKEKCKLLYQQ ARLHDROAAEMVLTIAGKGRETOMVAAT LKIGIALINGGRSTVQCKMLDYLKEKRDVGI POSLAGLMQSCSVULDDAFFERQKAKEGLOM VTEEGSGEK VLQDDEFTCDLFRETQLLCEGH NSPGPVLRTQTGANTTVAHISTYDYLLRVC ESISDPYWYTSGKDVIDEOGORNFSKAIQVA KQVFINTLTEYIQDCTCHOQGSLAHRSLWDAA VGFLHVFAHMOWKLSQDSSGELLKELMDLL KDMVVMLISMLEGNVNOTIGKGMVDML ESSNNVEMILKFFDMFLKLKDLTSSDTFKEYT PDGKGVYFRKDFHKAMESHHYTQSSTEFFLL SCAETDENETLDVEEFVKRFHEPAKDIGFNVV VLLTILESHMPNDTRUQFILLEAESVINYQ FLOREMGSAKRIERVYFEISESSKTQWKKY VKSSKGPIFDVVNEGGERKEMELFVNTGED TIFEMQLAAQISESDLNESSANKESSEKSREE QGRRMAFFSILLVRSALFALKNILLTILMSMLE LKSLKQMKKVKKMIVKDMVTAFFSSYVSI FMTLLHTVASVFREFFRICCLLLGGSLUSGV KGA KKKVAELLANMPDFTQBGEKKMELFVNTGED TIFEMQLAAQISESDLNESSANKESSEKSREE LEAALPSDLTDLKELFESDLLSDFLGJV LORG KERKAKEDKGKÇKLRQLHTHRYGEPEVPES, FWKIJIYQQKLLNYARNFNRYMEMPLVGE FROAKEREESPLSSSFROHED FRANKING AKKKVAELLANMPDFTQBGERKEMELTLINSMIL LKSLKRQMKKVKKMIVKDMVTAFFSSYVSI FMTLHTVASVFREFFRICCLLLGGSLUSGV KOA KKIKVAELLANMPDFTQBGERKGMELFILMSMLL LGGGQYKLIPHINPAGLSDLKGOPENGERGERKEMELFILSSULTSDFFLOH THYNTYSFCIIGTYCTLAVFANTSSPFREF THYNTYSFCIIGTYCTLAVFYNTOSFP NNYWDKFVKRKPMBACKPKSDSSSS,SAVLNSID VKYQMKLGOUTTONGSTYNTOKPPTVRIC PLHTVISFCIIGTYCTLAVFRICESPLATH NGRGGLUTTVGLGSTYWKMYPREFTVRF THYNTYMPHYVLESSCOFT SAVLNSID VKYQMWLGOUTTONGSTYMMYMPROTACH				ļ			J	
AEMCCTALNSEHMITLLGRILKHITNIN.I.GIDI GAWMKRLAYSEGDHI KARARGOMSEAL ILDEFTTLARDI YAPPILLIRGOYNRAKW KERNPEAEELFRAYAEVITWSKSHITKREE ONFOVONEINMISHITUTISKSMAKAYSD ERKKMKRKGORYSMOTSI.VAALKRIL PIQI NICAPODQELLALAKIRSI.KDTEDEVADUR. NIFLIQOKLEDPARWOMAI.YKDLPINRTDDT DPEKTYERVIDIANUF PILEOKSKKAVWIKLLSKQRKKAVVACI RMAPI.YMJPRIRANAVILLQOYEKSWIETEB YFEDKLEDLAKFGAEPPEDEGTRKVOPILH LILLISKTATITEKCKLEEDI.YMAYADIMAKK CHOBEDDDGEEVKSPEREMERKGKLI YQQ ARALHDROAABWI.QTISASKGETGPMVAAT LKJGIALIMGGISTYOVQOKMOTYLKKKDVGI FQSLAGLIMGGCSVI.DLNAFERQINKAEGLOR VTEEDSGEKVI.QDDEFTCDLFREI,DLLECH NSDFQNYLRTQTGNNTTVNIIISTVDYLLRVQ ESISDFWYYSGKOVIDEOQGRRFSKAIQVA KQVFNTLTEYIQOPCTONQOSLAHSRI.WDAA VGFLHVAHMOMAIS.OPSSGIELLKEIMDL KDMVVMLISMLEGNVVNOTIGKQMVDML ESINNVEMILKFPDBFLKI.ELDISDTFREYY PDOKGVIFKROPHKLIKDLTSSDTFREYY PDOKGVIFKROPHKLIKDLTSSDTFREYY PDOKGVIFKROPHKLIKLDLTSSDTFREYY PDOKGVIFKROPHKLIKDLTSSDTFREYY PDOKGVIFKROPHKLIKDLTSSDTFREYY POKGVIFKROPHKLIKDLTSSDTFREYY POKGVIFKROPHKLIKDLTSSDTFREYY POKGVIFKROPHKLIKDLTSSDTFREYY POKGVIFKROPHKLIKDLTSSTTWEKP VKESKROPIFDVNRGGEREKMEI.FYNNCEDE TIFEMQLAAQISESDLESSANKESSTREE QGPRMAFFSILTVRSALFALRYNILTLMRMI. LKSLKKOMKKVKKMITVSDMVTAFSSYNNS FMILLHFVASYFRGFFRICSLLLGGSLVGGA KIKKVAELLANMPOPTOOPVRGGEREKMEI.FYNNCEDE TIFEMQLAAQISESDLESSANKESSEKREE QGPRMAFFSILTVRSALFALRYNILTLMRMI. LKSLKKOMKKVKKMITVSDMVTAFSSYNNS FMILLHFVASYFRGFFRICSLLLGGSLVGGA KKIKVAELLANMPOPTOOPVRGGEREKMEI.FYNNCEDE GGGYKLIPHNPNGLSJDLISSDAYNESSEKREE QGPRMAFFSILTVRSALFALRYNILTLMRMI. LKSLKKOMKKVKKMITVSDMVTAFSSYNNS FMILLHFVASYFRGFFRICSLLLGGSLVGGA KKIKVAELLANMPOPTOOPVRGDGERGERKI LEAALPSEDLTDLRELTESSDLJSDFGIDLK GROUPT TO THE TOTAL THE TOT	Ì							
GAWMKRLAVFSQRINKVRPQLLKTHEIPLIN EKLKKKAATVVSEGDHIKABARGDMSEAEL LILDEFITLARDLYAFYPLLIRFGDYNRAKW KEPNPEAELFRAWAEVIFWSKSKINFKREE QNFVVQNEINMSELITDITSKMSKAAVSDE BEKKMRKGGRYSMOTSLIVAALKALLPIGI NICAPGDQELALAKNRFSLKDTEDEVEDUR. NIHLQCKLEPPARWQMALYKLDENRTDDT: DPEKTVERVIDIANVLFHLEQKSKRVGRRIY CLVEHPQBSKKAVWHKLLSKQRKRAVVACI RMAPLYNIPRIRAWNLFLQGYEKSWIETEE YFEDKLEDLAKFGAEPPEEDGTKRVDPLH LILLFSKALTEKCKLEEDFLYMYADIMAK. CHDEEDDDGEEEVKSPEEKEMERQKLLYQQ ARLHDBGAAEMYLQTISASKGETGFMVVAAT LKLGIAILNGGNSTVQQKMLDYLKEKDDVG FQSLAGLMQSCSVLDLAFFEQNKAEGLOM VTEEGSGEKVLQDDEFTCDLFRFLQLLCEGH NSDFQWYLRTGTGRNTTVMIISTVDYLLRVQ ESISDFYWYYSGKDVIDEQGQRNFSKAIQVA KQVFNTLTEYIQGPCTGNQSLAHSRWDAA VGFLHVFAHMQMKLSQDSSQIELLKELMDL KDMVVALLSMLEGNVNOTIGKGMVDML ESSNNVEMILKFFDMFLRLLDLTSSDTFKEYY PDCKGVFKRDPHKAMESHKHYTOSTEFFLL SCAETDENETLDVEEFVKRFHEPAKDIGFNV, VLITNLSSHMPDITELQTFLLEASEVLNYPQ FLORGWYKRDPHKAMESHKHYTOSTEFFLL SCAETDENETLDVEEFVKRFHEPAKDIGFNV, VLITNLSSHMPDITELQTFLLASEVLNYPQ FLORGWYKRDPHKAMESHKHYTOSTEFFLL SCAETDENETLDVEEFVKRFHEPAKDIGFNV, VLITNLSSHMPDITELQTFLLASEVLNYPQ FLORGWYKRDPHKAMESHKHYTOSTEFFLL SCAETDENETLDVEEFVKRFHEPAKDIGFNV, VLITNLSSHMPDITELQTFLLASEVLNYPQ FLORGMAFFSILTVRSALFALRYNILTLARMSL LKSLKKQMKKVKKMVKMVTAFFSSTWS FMILHFVASVFRGFFRICSLLLGGSLVBGA KKKVAELLANMPPOTQOFTKOBGEGERK LEAALPSEDLTDLKELTEESDLLSDFFGLDX FREMQLAAGSSBDLNESANRESSEKRREI QGRWAFFSILTVRSALFALRYNLITLARMSL LKSLKCQMKKVKKMVKNVTAFFSSTWS FMILHFVASVFRGFFRICSLLLGGSLVBGA KKKVAELLANMPPOTQOFTKOBGEGERK LEAALPSEDLTDLKELTEESDLLSDFFGLDX GKYGKYLHPINPNAGLSDLLSDLSSDFFOLDX GKYGKPHPPTVRGCBGGERK LEAALPSEDLTDLKELTEESDLLSDFFOLDX GKYGKPHPPTVRGCBGGRKK KKRVAELLANMPPTQOFTKOBGEGERK LEAALPSEDLTDLKRLTESSSCYMEPTVRII PLHTVISFFCIIGTYCLKYPLVIFKREKEVARK LEFDGI YITGPSEDDIKGQWRRLVINTOSTPF NYWDYSKYRKPWHDYVLEESSCYMEPTVRII PLHTVISFFCIIGTYCLKYPLVIFKREKEVARK LEFDGI YITGPSEDDIKGQWRRLVINTOSTPF NYWDYSKYRENDYNTMSVPHYNGRSELLG GGQYKLIPHNPNAGLSDLLSDLSSTYNTMSVYNTMSVYL GHYNNFFEAAHLLDLAMGFKTRTILSSVTH NGKQLVITVGLLAVVYVYLTYVVARNFFRRF YNSSELOGTPDMCCDDMCTYMFTRIPSLFF YNSGLODTPDMCCDDMCTYMFTRIPSLFF TOTULCCSGWSAVAPSRLTAALFS*AQAVCL SPRSWOVRNWPHPANNCIFC	١							AEMCCTALNSEHMNTLLGNILKIIYNNLGIDE
EKLKKKAATVVSÉEDHLKAEARGÖMSEK LILDEFTTLADI/YAFYPLIRGÖYNRAKW KEPNPEAEELIRMVAEVFIYWSKSINIKRAAVSDO ERKKMKRGDRYSMOTSLIVAALKRILPIGI NICAPODELIALAKNIPSLIVESWISKSANAVSDO ERKKMKRGDRYSMOTSLIVAALKRILPIGI NICAPODELIALAKNIPSLIVESVENSKENAVSDO PEKTVERVLDIANVILFILEGSVENSKEVACIO. RMAPLYNLPHRAVNILFLOGYEKSWIPTEE YFEDKLEDLAKPGAEPPEDEGTKRVDPLH LILLFSRTALTEKCKLEEDFLYMAYADIMAK. CHDEEDDOGEEVKSFEKSEMERGKLLYQQ ARLHDROAAEMVLQTISASKGETGEMVAA. CHDEEDDOGEEVKSFEKSEMERGKLLYQQ ARLHDROAAEMVLQTISASKGETGEMVAA. LILLGIALNGGNSTVQOKMLDYLKEKKDVGI FQSLAGLMQSCSVLDLAFERQNKLAGLGOM VTEEGSGEKVLQDDEFTCDLFREQULCEGH NSDEQNYLRTQTGNNTTVNIIISTVDYLLRGV ESISDYWYYSKENDVIDEQGGRNFSKAIQVA KQVFNTLTEYIQGPCTGNQSLAHSRUWAV CESISDYWYYSKENDVIDEQGGRNFSKAIQVA KQVFNTLTEYIQGPCTGNQSLAHSRUWAV LESSNNVEMILKFPOMELKLENDL KDMVVALLSMLEGNVVNGTIGKQMVDMU LESSNNVEMILKFPOMELKLENDL KDMVVALLSMLEGNVVNGTIGKQMVDMU LESSNNVEMILKFPOMELKLENDL KDMVVVDLESWERFFREMERHENTYTQSETETEL BERNVEMILKFPOMELKLENDLYSSDTFKEYI PDGKGVFKRDFHKAMESHKHYTQSETETEL LEGLEIMGSAKRERVFYEESESSTROWEKPQ VKESKRQFIFDVVNEGGEREKMELFVYTQSETETEL LEGLEIMGSAKRERVFYEESESSRTOWEKPQ VKESKRQFIFDVVNEGGEREKMELFVTVPQ FLGEIEMGSAKREVFYEESESSRTOWEKPQ VKESKRQFIFDVVNEGGEREKMELFVTVPQ FLGEIEMGSAKKKMITVKDMVTAFFSSYWS KMILLHVASVPRGFFRIICSLLLGGSLVGGA KKIKVAELLANMPPTQDEVRGDEEGERK LEAALPSEDLTDLKELTEESDLLSDIFGLDLKI EGGGYKLJEHNPAGLSDLANGTANGSEPEREE VARK LEAALPSEDLTDLKELTEESDLLSDIFGLDLKI EGGGYKLJEHNPAGLSDLANGTANGSEPEREE VARK LEAALPSEDLTDLKELTEESDLLSDIFGLDLKI EGGGYKLJEHNPAGLSDLANGTANGSEPEREE VARK LEFDGLYTTEQPSEDDIKGQWDRLVINTIGSFP NNYWBSYKRKVAMDKYGEFYGRDRISLING GHYNNFFFAAHLDLAMGFFTLKTLSSVTH NGKQLVLITYGLLAVVVYLYTVVAFNFREY VKSGBOTPDMKCODMLTCVTMFHMYVGV RAGGGIGBEIDPAGDENDQRGVYEDMETKC FLOGIGNDYFDTVPRIGFETHTLGENLLANVY VKLIALIGGLIUDAGELERQGERGVKEMYFYVE VILLALIGGLIUDAGELERQGERGVKEMYFYVE VILLALIGGLIUDAGELERQGERGVKEMYFYVE VILLALIGGLIUDAGELERQGERGVKEMYFYVE LEGLENDSDPPALASQSAGITGVTRTPSLFF DTVILLCCSGWSAVAPSBLITAALFS*AQAVC SISPSWDYRKWPPHANFICTCFGDESILAML	- [[1		1	
KEPNPEAGELPRMAEVFYWSKSINKEKSKSKAAVSDO ERKKMKRGDRYSMOTSLIVAALKRILPIGI NICAPODGELIALAKNEFSLEVEKSKSKAAVSDO ERKKMKRKGDRYSMOTSLIVAALKRILPIGI NICAPODGELIALAKNEFSLEVEBEVRDIKI NIHLOGKLEDPAIRWQMALYKDLPNRTDDT: DPEKTYPRYDLDIAVILPHLEQKSKRVGRYIV CLVEHPQRSKKAWHKLLSKORKRAVVACI RMAPLYNLPKHRAVNLFLQGYBESSWETEST YFEDKLEDLAKVAWHKLLSKORKRAVVACI RMAPLYNLPKHRAVNLFLQGYBESSWETEST YFEDKLEDLAKOAPPEPEDEGTKRYDPLH LILLFSRTALTEKCKLEEDFLYMAYADIMAK CHDEEDDDGEESVKSFEEKBRGKLLYQQ ARLHDRGAAEMVLQTISASKGETGMYAADIMAK CHDEEDDDGEESVKSFEEKBRGKLLYQQ ARLHDRGAAEMVLQTISASKGETGMYAADIMAK CHDEEDDDGEESVKSFEEKBRGLLCEGH NSDFQNYLRTQTGNNTTVNIIISTVDYLLTAC FOSLAGLMOSCSVLDLNAFERQNKAEGLGM VTEEGSGEKVLQDDEFTCDLFEQLLCEGH NSDFQNYLRTQTGNNTTVNIIISTVDYLLTAC ESISDFYWYYSKORDVDGOGQRNSKARQVA KQVFNTLTEVIQGPCTGNQSSLAHSRLWDAA VGFLHVFAHMQMKLSQDSSQIELKERIMDLA KDMVVALLSMLEGNVVNGTIGKQMVDML ESSNNVBMLKFFDMFLKLKDLTSSDTFSLY PDGKGVFKROPHEAMSSHHYTQSSTEFFLL SCAETDENSTLDYEEFVKRHEPAADIGTNV VLLTINLSEHMMPDITLQTTLELASVLNYPQ FLGRIEMGSAKRIERVYTEISESSRTQWEKY VKESKROFFIDVYNEGGEKEKMELFVNPCED TIFEMQLAAQISESDLNERSANKESSEKEREN VKESKROFFIDVYNEGGEKEKMELFVNPCED TIFEMQLAAQISESDLNERSANKESSEKEREN LEAALPSBOLTIN JEKTLEFSDL STOFLOLD LEGGGYKLIPHNNAGLSDLMSNPVPMPEVQE KPQEGYAKEEKEKETNSEPEKABOGDOG KERKAREDKOKQKLRQHTHRYNGTSDFALL LEAALPSBOLTIN JEKTLEFSDL STOFLOLD LEGGGYKLIPHNNAGLSDLMSNPVPMPEVQE KPQEQKAKEEKEKETNSEPEKABOGDOG KERKAREDKOKQKLRQHTHRYSSSENA KKIVABLLANMPPOTODEVRODEEGERKK LEAALPSBOLTIN JEKTLEFSDL STOFLOLD LEGGGYKLIPHNNAGLSDLMSNPVPMPEVQE KPQEQKAKEEKEETKSEPEKABOGDOGE KERKAREDKOKQKLRQHTHRYSTSVVORGKLPTRSSSENA KVTSLDSSSHRIMAYHYVLEESSGYMEPTYVA AFANNELLFYKVSTSSVVORGKLPTRSSSENA KVTSLDSSSHRIMAYHYVLEESSGYMEPTYVA AFANNELLFYKVSTSSVVORGKLPTRSSSENA KVTSLDSSSHRIMAYHYVLEESSGYMEPTYVA HVNNFFPRAHLLDIAMGFKTLRTLISSVENA KVTSLDSSSHRIMAYHYVLEESSGYMEPTYVA NGKQLVLTVGLALAVVYLYTYVARNPFRYK NGKQLVLTVGLALAVVYLYTYVARNPFRYK NGKGGIGDEIDPAGGEROKECKETKSPFREABOGDOE KERKAREDKOKGKLRQHTHRYSSENA LEFOGLYTTEQFSEDDIKGGWORLVUNTQSPP NNYWDKFVKRKVMOKYGGEPOKEPVEN NGKGGGGDEEDPAGGEPOKENAM LEFOGLYTTEGESTLAMA LEFOGLYTTEGSTENDALTALFS*AQAVCL SLEENDDDPALAGOSAGITGVTRTPS	1)	EKLKKKAATVVSEEDHLKAEARGDMSEAEL
QNFVVQNEINNMSFLITDTKSKMSKAAVSDE ERKKMARKGDVSMQTSILVAKRLI PIGI NICAPGOGELIALAKNEFSLKDTEDEVRDIKE NIHLQGKLEDPARWQMALYKDLPNTDDT DPEKTVERVLDIANVLPHLEQKSKRVORRIV CLVEHPQRSKAVWHLLISKQRKRVVVAC (RMAPLYNLPHRAVNLFLQGVESSWIETEE YFEDKLEDLAKPGAEPPEDEGTKRVDH- LILLFSRTALTEKCKLEEDLY-MAYADIMAK CHDEEDDDGEEVKSFEKEMEKQKLLYQQ ARLHDRGAAEMVLQTISASKGETOMYVAAT LKLGJALINGGNSTVQQKMLDYLKEKKDVG FOSLAGLMQSCSVLDLNAFERGMVAAFGLGM VTEGGSGEKVLODDEFTCDLFRRLQLLCEGH NSDFQNYLRTQTGNNTTVNIIISTVDYLLRW ESISDFYWYYSGKDVIDEQGQRFSKAIQVA KQVFNTLTEYIQGCTGNQQSLAHSRLWDIA VGFLHVFAHMQMKLSQDSSQIELKELMDLA KDMVVALLSMLEGNVVNGTIGKQMVDMU ESISNVEMILKFPDMFLKLKDLTSSDTFKEYI PDGKGVFKRDFHKAMESHKHYTQSSTEFLE LEGGLEIMGSAKREVYFSESESSRTQWEKP FLGEIEMGSAKREVYFSESESSRTQWEKP FLGEIEMGSAKREVYFSESESSRTQWEKP VKESKROFFDVVNEGGEKEKMELFVNCGD TIFEMQLAAQISESDLNERSANKEESEKERPE QQFRMAFSILTVRSALFALRYNTLINMMLS LKSLKQMKKVKKMTVKDMVTAFFSSVYBG VKESKROFFDVVNEGGEKEKMELFVNCCDB TIFEMQLAAQISESDLNERSANKEESEKERPE GQGYKLIPHNPNAGLSDIMSNPVPMPEVC KFGGQYKLIPHNPNAGLSDIMSNPVPMPEVC KFGGQYKLIPHNPNAGLSDIMSNPVPMPEVC KFGGQYKLIPHNPNAGLSDIMSNPVPMPEVC KFGGQYKLIPHNPNAGLSDIMSNPVPMPEVC KFGGGYKLIPHNPNAGLSDIMSNPVPMPEVFINI LEAALPSBLITULKELTEESDLINGELDLE KKIKVAELLANMPDPTQDEVRGDGEEGERKRE LEAALPSBLITULKELTEESDLINGELDLE KKIKVAELLANMPDPTQDEVRGDGEEGER KEEKAREDKGKQKLRQLHTHRYGGPEVPFS FWKKILYQQKLUNYANNFVNMMALAFV FAANNFLLFYKVSTSSVVGGKELPTRSSSENA KVTSLDSSRRIJAVHYANFNFYNMRALAFV FAANNFLLFYKVSTSSVVGGKELPTRSSSENA KVTSLDSSRRIJAVHYVLEESSGYMEPIVVRI PILHTVISFFCIGGYYCLKYPLVIFKREKEVARK LEFOGLYTTEGFSPLOMKCQ MDKAALDFSDAREKKFKDSSSLSAVLNSID VKYQMWKLGVVTTONSFLYLAWYMTMSVLYV GHYNNFFFAAHLLDIAMGFKTLRTILSVTHY NGKQLALTVYQLTYVAPNFFYFY NGKGQUTTOVALTAVTYLYTVAPNFFFKF NGKSEDGDTPDMKCDDMLTCYMFHMYVGY RAGGIGDEEDBPAGDEYEYRITTFFFYT NGKQLALTVYQLTYVAPNFFYFYR NGKGQUTTOVALTAVYYLYTVAPNFFFKF NGKSEDGDTPDMKCDDMLTCYMFHMYVGY RAGGIGDEEDBPAGDEYEYRITTFFFYT NGKGLUTTVOLLAVVYLYTVAPNFFYFYR NGKGLUTTVOLLAVVYLYTVAPNFFFKF NGKSEDGDTPDMKCDDMLTCYMFHMYVCR FFPAGDCTRKQYFEQELY SLEELNSDDPPALASGSGGTGVTRTPSLFF PTULLCCSGWSAVAPSSLITAALFS*AQVCL SLFRSWDYRKWPPHANFICTCFGDESLAWCL SLFRSWDYRKWPPHANFICTCFGDESLA	ļ				!			LILDEFTTLARDLYAFYPLLIRFGDYNRAKWL
ERKKMKRGDRYSMOTSLIVAALERLIPIED INICAPGOGELIALAKNRISLADEVARDIN NIHLQGKLEDPARWQMALYKDLPNRTDEVRDIR NIHLQGKLEDPARWQMALYKDLPNRTDEVBUR NIHLQGKLEDPARWQMALYKDLPNRTDEVBUR PERTVERVLIDLANVLHELGKYKRVGPRIN CLVEHPQRSKKAVWHKLLSKQRKRAVVACI RMAPLYNLPHRAVANLFLQOYEKSWETEEB YFERKLEDLARGACPFPEDEGTKRVDPLH LILLIFSRTALTEKCKLEEDFLYMAYADIAMAK CHDEEDDDGEEGKKSPEEKEMEKQKLLYQQ ARLHDRGAAEMVLQTISASKGETGMVAACI LKLGIALINGGNSTVQQKMLDYLKEKKDVGI FQSLAGLMQSCSVLDLNAFERQNKLAEGLLCHGM VTEEGSGEKVLODDEFTCDLFREQLLCEGM NSDPQNYLRTQTGNNTTVNHISTVDYLLKYC ESISDPYWYYSGKDVDIEQGQRNFSKAIQVA KQVFNTLTEVIQGPCTGNQSLAHSRLWAV VESHLVAHMQMKLSQDSSOJELLKELMDLA KDMVVALLSMLEGNVVNGTIGKQMVDMLA LESNNVEMILKFFDMFLKLKDLTSSDTFREYI PDGKOVFKRFFHKAMESHRHYTQSETEFLL SCAETDENSTLDYEEFVKRFHEPAKDIGFNV VILLINLSEMMPMDTILQTTLELASVLNYPQ FLGRIEIMGSAKRIERVYFEISESSRTQWEKS VYKESKRQFFFDVYNEGGEEKKMELFVNFCED TIFEMQLAAQISESDLNERSANKESSEKERE QOPRMAFFSILTVRSALFLARLYNLTLMRMLS LKSLKQMKKVKKMTVVKDMVTAFFSSVVSIS FMTILHFVASVPGFFRIEGLIJGSVLSGA KKIKVAELLANMPDPTQDEVRGDGEEGERKR LEAALPSBLDTLARLETEESDLAFGLJEGLUG KCEKAKEDKGKQKLRQLHTHRYQEFEVPESS FWKRIIAYQQKLLNYFARRNFYNMRMLALFY AFANNELLFYKVSTSSVVEGKELFVNSSSENA KVTSLDSSSHRIANHYVLEESSGYMEPTVYAF AFANNELLFYKVSTSSVVEGKELFVNSSSENA KVTSLDSSSHRIANHYVLEESSGYMEPTVYAF AFANNELLFYKVSTSSVVEGKELFVNSSSENA KVTSLDSSSHRIANHYVLEESSGYMEPTVYAF AFANNELLFYKVSTSSVVEGKELFTRSSSENA KVTSLDSSSHRIANHYVLEESSGYMEPTVYAF AFANNELLFYKVSTSSVVEGKELFTRSSSENA KVTSLDSSSHRIANHYVLEESSGYMEPTVYAF AFANNELLFYKVSTSSVVEGKELFTRSSSENA KVTSLDSSSHRIANHYVLEESSGYMEPTVYAF NGKQLJAVVYATIVVARNFYRKP NGKQLJAVVYATIVVARNFYRKP NGKQLJAVVYATIVVARNFYRKP NGKQLJAVVYATIVVARNFYRKP NGKQLJAVVYATIVARNFYNKNSUL GHYNNFFRAHLLDIAMGFKTLRTLISSUENA KVTSLDSSSHRIANHYVLEESSGYMEPTVYAN NGKQLJAVVYATIVARNFYNKY NGKGLDALAVVYATIVARNFYNKY NGKGLDALAVVYATIVARNFYNKY NGKGLDALAVVYATIVARNFYNKY NGKGLDALAVVYATIVARNFYNKY NGKGLDALAVVYATIVARNFYNKY NGKGLDALAVVYATIVARNFYNKY NGKGLDALAVVYATIVARNFYNKY NGKGLDALAVVYATIVARNFYNKY NGKGLDALAVVYATIVARNFYNKY NGKGLDALAVVYATIVARNFYNKY NGKGLALAVVYATIVARNFYNKY NGKGLALAVVYATIVARNFYNKY NGKGLALAUALATALTH NGKGLALAVVYATIVARNFYNKYNGALLALTH NGKGLAL		•						KEPNPEAEELFRMVAEVFIYWSKSHNFKREE
NICAGEDGELIALAKNRISLKOTEGEVRIDIK NIHLGGKLEDPARROWANU, PINTEDDT DPEKTYER VLDIANVI, FILEGKSKR VGRIN CLVEHPORSKA, VWHKLEN, KORKRA VVAGC RMAPL YNLPRHRAVNI, FLQGYEKSWIETEER YFEDKLIEDLAKRGAEPPEEDEGTKREVDI, HI LILLFSSTALTEKCKLEEDPI, YMA YADIMAK. CHDEEDDDGEEVKSFEKEMEKQKLI, YQQ ARLHDRGAAGMVL, GYISASKGETGFMVAAT LKI, GIALINGGISTVQ, KMLDVL, KEKRDVGI FQSLAGLMNG, GSVL, DLNAFRONKAEGLGM VTEGSGEKVL, QDDEFTCDLFREI, QLLCEGH NSDFONYLRTGTGNNTTVINSTVDYLLRVC ESISDFYWYYSGKDVIDEQGRNFSKALQVA KQVFNTLETYQOPCTGNQSLAHSRL WDAN VGFLHVFAHMQMKLSQDSSQIELKELMDILL KDMVVMLLSMLEGNVVNGTIGKQMVDML ESSNNVEMILKFFDMFLKLKTSSTFFKLY PDGKGVUFKRDFHKAMESHKHYTQSETFFLL SCAETIDENETLDVEEPVKRFHEPAKDIGFNV- VLLTHLSEHMFNDTRLQTIFLELAESVLNYFQ FLGREIMGSAKRIERVYFEISSTFOWENCH WESSKROFFTDVVNEUGEREKKMET, FVNFCD TFEMQLAAQISESDLNERSSENTQWEKKQ VKESKROFFTDVVNEUGEREKKMET, FVNFCD TFEMQLAAQISESDLNERSSENTQWEKKQ VKESKROFFTDVVNEUGEREKKMET, FVNFCD TFEMQLAAQISESDLNERSSENTQWEKKQ VKESKROFFTDVVNEUGEREKKMET, FVNFCD TFEMQLAAQISESDLNERSSENTQWEKKQ VKESKROFFTDVVNEUGEREKKMET, FVNFCD TFEMQLAAQISESDLNERSSENTQWEKKQ VKESKROFFTDVVNEUGEREKKMET, FVNFCD TFEMQLAAQISESDLNERSSENTQWEKKQ VKESKROFFTDVVNEUGEREKKMET, FVNFCD TFEMQLAAQISESDLNERSNEESKEREPER QOPRMAFFSILTVRSALFALRYNILTLAWMLS LKSLKKQMKKVKKMTVKDMVTAFFSSYWS FMILLHEVASVFRGFFRICSLLLGGSLVEGA KKIKVABLLANMPDPTQDEVRGDGEGRKI LEAALPSEDLTDLKELTEESDLLSDIFGDLKI EGGGVKLIPHNPNAGL,SDLMSNPVPMPVQEC KFQEKAKEEKEKEETISTSPFEKAEGEDGE KEEKAKEDKGRQKLRQLHTTRYSEPPVENS FWKKILTYQVGKLLNYFARNFYNMRMLALFV AANIFILLFKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRILAHTVVLGERFYGRDRISELLG MDKAALDFSDAREKKRKKKDSSLSAVINSID VKYQMWKLGVVTUNSFP, TVAWYMTMSVL, GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVTLVYAFNFRKFK YNKSEDGDTPDMKCDDMLTCVMFHRINYOV RAGGGIGDEEPAGDEVFRUIPDTFFFFV VULLAUQLLUDAGELRDQOEQVKEDMBTKC FICGIGNDYFDTVPHGFETHTLQERNLANVILF FLMYLINKDETHTIOQESPYWKDMYGRCWE FFRAGDCFRKQVEDQLN AGLELLINSDDPPALASQSAGGITGVTRTPSLFF DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRWPPHAAPICTROESLAML SLPRSWDYRRWPPHAAPICTROESLAML SLPRSWDYRRWPPHAAPICTROESLAML SLPRSWDYRRWPPHAAPICTERGESLAML	1				<u>'</u>		:	QNFVVQNEINNMSFLITDTKSKMSKAAVSDQ
NIHLQGKLEDPARWQMALYKDLPNITDTD DPEKTVERVDIANUFHLEQKSKRVGRRIJY CLVEHPQRSKKAVWHIKLISKQRKRAVVACI RMAPLYNLPRIRAVNIELLGKSKRVGRRIJY CLVEHPQRSKKAVWHIKLISKQRKRAVVACI RMAPLYNLPRIRAVNIELLGKSKRVGRRIJY FORKLIEDLAKPGAEPPEEDEGTKRVDPH LILLFSRTALTEKCKLEEDFLYWAYADIMAK CHDEEDDDGEECWSFEFEERKGKLLVQQ ARLHDRGAAEMVLQTISASKGETGFMVAAT LKLGIALINGGNSTVQQKMLDVLKEKKDVGI FOSLAGLMQSCSVLDLNAFERQNKAEGLGM VTEEGSGEKVLQDDEFTCGNLDTYRLISTVDYLLRVC ESISDFYWYYSGKDVIDEQGQRNFSKAIQVA KQVFINLTEYTQOPCTGNQGSLAHSRLWDAA VGFLHVFAHMQMKLSQDSSQIELLKEMDL KDMVVMLLSHDENVONTGTGKQMYDML ESSNNVEMILKFFDMFLKLKDLTSSDTFKEYI PPOGKGVIFKRDFHKAMESHTYQESTEFLL SCAETDENETLDVEEFVKRFHEPAKDIGFNV VLLTALSEHMPDTRITQTFLAESVLNYVQ FLGRIEMGSAKRIERVYFEISESSKTQWEKPQ VKESKRQFHDVVNEGGEKEKMELFVNFCDD TIFEMQLAAQISESDLNERSANKEESSEKREWEL QGPRMAFFSLTVRSALFALRYYNLTHRMLS LKSLKKQMKKVKKMTVKDMVTAFFSSYWS FMTLLHFVASVFRGFFRICSLLLGGSLVGA KKIKVABLLANMPDPTQDEVRGDGEGEGRK LEAALPSEDLTDLKELTEEDLLSDIFGCDLKI EGGGYKLIPHPRNAGLSDLAMPYNDRAMPYPMPEVOE KFQEQKAKEEKEEKEKEFKSEPEKAEGEDGE KEEKAAEDKOKQKLRQLHTTRYGCEFPVESS FWKKILAYQUKALNYFARMYNMRMLALFV AFANNFLLFVKVSTSSVVEGGELFTRSSSENA KVTSLDSSSHILAVHYVLEGPSVESSENA KVTSLDSSSHILAVHYVLEGPSVESSENA KVTSLDSSSHILAVHYVLEGPSVESSENA KVTSLDSSSHILAVHYVLEGPSVESSENA LEFGOLNTIFOFSEDDLKGOWDRLVINTOSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKKPKKDSSLSAVLNSID VKYQMWKLOVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDLAMGFKTLRTILSSVTH NGKQLUTVGLAVVVYLVAFNFRRKF YNKSEDGDTPDMKCDDMLTCYMFHINYOV RAGGGIGDEEPAGDEVFRUFTIFFTYV VILLAIQGLIDAFGELROQQEQVKEDMETTCC FICCIONDY FDTVPHGFETTHTQUENLANVILF FLMYLINKDETEHTQGESYVWKMYQERCWE FFRODCFRKQVEDDIN CHECKTORY PLAMPLINKDETEHTGGESYVWKMYQERCWE FFRODCFRKQVEDDIN AGGGIGDDEEDPAGDEVFRUFTEFTFFYV VILLAIQGLIDAFGELROQGEQVKEMBETTCC FICCIONDY FDTVPHGFETTHTQUENLANVILF FLMYLINKDETEHTGGESYVWKMYQERCWE FFRODCFRKQVEDDIN CHECKTORY PLAMPLINKDETEHTGGESTALMALLSSTAN TALLFSTAN TA							ĺ ' !	ERKKMKRKGDRYSMQTSLIVAALKRLLPIGL
DPEKTYERVLDIANVLFHLEQKSKRVQRINY CLVEHPQRSKAVWHKLISQRRRAVVAG, RMAPLYNLPRIRAVNLFLQGYEKSWIETEE YFEDKLIEDLAKPGAEPPEEDEGTKRVDPLH LILLFSRTALTEKCKLEEDLYMAYADIMAK. CHDEEDDDGEEVKSFEREMEKQKUTQQ ARLHDRGAAEMVLQTISAGTGMVAAT LKIGIALINGGNSTVQOKMLDVLKSKDVGI FQSLAGLMQGSVLDLNAFRQNKAEGLGM VTEEGSGEKVLQDDEFTCDLFRFLQLLCEGH NSDFQNYLRTQTGNNTTVNIIISTVDYLLRVG ESISDFYWYSGKDVIDEGQGNFSKAIQVA KQVFNYLTEYIQQPCTGNQQSLAISRLWDAN VGFLHVFAHMQMKLSQDSGLIKELMDDL KDMVVMLLSMLEGNVYNGTIGKQMVDML ESSNNVEMILFFDMFLKLNTSSDTFREYP PGGGVIFKRDFHKAMESHKHYTQSETFFLL SCAETDENETLDYEEFVKRFHEPAKDIGFNV VLLITALSEHMPNDTRLQTIT-ELAESVLNYFQ FLGREIMGSAKRIERVYFEISESSKTQWEKQ VEESKQPIFDVVNEGGEKMELFVNCEGE TIFEMQLAQGISESDLNERSANKESSEKERE QOPMAFFSILTYRSALFARVIT-LLAMEN, LKSLKKQMKKVKMTVKDMVTAFFSSVWS FMTLLHFVASVFRFFICISLLIGGSLVEGA KKIKVAELLANMPDPTQDEVRODGEEGGRK LEAALPSEDLTDLKELTESDLLSDIFGLDLKI EGGGYKLIPHNNAGLSDLWSNYPMPEVQE KFQEGKAKEEKEEKEFTKSFPEKAGEDGE KEEKAKEDKGKQKLRQLHTRRYGGPEVPESS KEEKAKEDKGKQKLRQLHTRRYGGPEVPESS KEEKAKEDKGKQKLRQLHTRRYGGPEVPESS FWKKIIAYQQKLLNYFARRFYYMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRILAVHYVLESSGYMEDTVKII PILHTVISFFCIIGYYCLKYPLVIFKREKEVARR LEPFOLNTTEGPSEDDLKGQWDRLVNTQSFP NNYWDKFVKRKVMDKYGEFYGRDISELIG MDKAALDFSDAAREKKFKKDSSLSAVLNSID MDKAALDFSDAAREKKFKKDSSLSAVLNSID MDKAALDFSDAAREKKFKKDSSLSAVLNSID MDKAALDFSDAAREKKFKKDSSLSAVLNSID MDKAALDFSDAAREKKFKKDSSLSAVLNSID MDKAALDFSDAAREKKFKKDSSLSAVLNSID MDKAALDFSDAAREKKFKKDSSLSAVLNSID MDKAALDFSDAAREKKFKKDSSLSAVLNSID MDKAALDFSDAAREKKFKKDSSLSAVLNSID MDKAALDFSDAAREKKFKKDSSLSAVLNSID MDKAALDFSDAAREKKFKKDSSLSAVLNSID MDKAALDFSDAAREKKFKKDSSLSAVLNSID MDKAALDFSDAAREKKFKKDSSLSAVLNSID MDKAALDFSDAAREKKFKTDITFFFYU VILLAIQGLIDDAFGELROQOGVKEDDMFTKC FICGIONFYFDTVHGFETHTHQUFTFFYY VILLAIQGLIDDAFGELROQOGVKEDDMFTKC FICGIONFYFDTVHGFETHTHQUFTFFYY VILLAIQGLIDDAFGELROQOGVKEDDMFTKC FICGIONFYFDTVHGFETHTIQEBINLANVIL- FLMYLINKDETHTTGGESYVWKMYQERCWE FFPAGDCTROYSDOALLALSSTAAVLDLANKSPRFTATALFS*AQAVCL SLPRSWDYRKWPPHAANFCERESLSANL	- 1							NICAPGDQELIALAKNRFSLKDTEDEVRDIIRS
CU-VEHPQRSKKAVWHKLLSKORKRAVVACI RMAPL-YNLPRIRAVNLFLQOYEKSWIETEE YFEDKLEDLAKPGAEPPEEDEGTKRYDPLH LILLFSRTALTEKCKLEEDFLYMAY ADIMAK CHDEEDDDGEEVKSFFEKERKEKQKLLYQQ ARLHDRGAAEWYLQTISIASKGETGFMVAAT LKLGIALINGGNSTVQQKMLDVYLKEKKDVGI FOSLAGLMQSCSVLDLNAFERQNKAEGLGM VTEEOSGEKVLQDEFTCJACHGVLCCH NSDFQNYLRTOTGNITTVNIISTVDYYLRVC ESISDFYWYSGKDVITEGORFSKAQVA KQVFINTLTEYIQGPCTGNQQSLAHSRLWDAN VGFLHVPAHMQMKLSQDSSQIELKELMDL KDMVVMLLSMLEGNVVNGTIGKQMYDML ESSNNVEMILKFFDMFLKLKDLTSSDTFKEYI PPGKGVIFKRDFHKAMESHYTQSETFELL SCAETDENETLDVEEFVKRFHEPAKDIGFNV VLLTHLSEHMPDTRLQTLEASVLNYPQ FLGRIEMGSAKRIERYYFEISESSKTQWEKR VKESKRQFHTDVVNECGEKEKMELFVNFCED TIFEMQLAAQISESDLNERSANKEESEKERFEI QQFRNAFFSLTVRSALFALRYNILTIKRMLS LKSKKQMKKVKKMTVKDMVTAFFSSYWS FMTLLHFVASVFRGFFRICSLLLGGSLVEGA KKIKVABLLAMMPDPTQDEFORGEGEKEK LEAALPSEDLTDLKELTEEDDLLSDIFGLDLKI EGGGYKLIPHIPPAGLSDLKOFGGEGEGKEK KKIKVABLLAMMPDPTQDEFORGGEGEGKEK LEAALPSEDLTDLKELTEEDDLLSDIFGLDLKI EGGGYKLIPHIPPAGLSDLKOFGGEGEGKEK KKIKVABLLAMMPDPTQDEFORGEGEGEKEK FWKKILAYQQKLLNYFARNFYNMRMLALFV AFANFELLFYKVSTSSVVEGKELFTSSENA KVTSLDSSSHRILAVIFVLEGFGEGEGEKEK FWKKILAYQQKLLNYFARNFYNMRMLALFV AFANFELLFYKSTSSVVEGKELFTSSENA KVTSLDSSSHRILAVIFVLEGFSEDDIKSOFMETVRIL FILHTVISFFCIIGYYCLKYPLJVIFKREKEVARK LEFFGLNTTEGPSEDDIKGOWDRLVINTOSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAAREKKFKNSDSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVJ GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLAVVYVLYVAFNFREKF YNKSEDGDTPDMKCDDMLTCYMFHINYOV RAGGGIGDEEPAGDEVFURIFEDTFFFYV VILLAIQGLIDDAFGELROQQEQVEKDMETTKC FICGIGNDYFDTYPHGFETTHJQEENLANVIL FLMYLINKDETEHTGGESYVWKMYQERCWE FFRAGOCFRKVEDDALNTAFS*QAVCL SLPRSWDYRWFPHANFOLGDESDLAML FLANFINGERDPADDEFTKGFTFFTFTY VILLAIQGLIDDAFGELROQGEQVEKDMETKC FICGIGNDYFDTYPHGFETTHJQEENLANVIL FLMYLINKDETEHTGGESYVWKMYQERCWE FFRAGOCFRKVEDDALN FFRAMFUR PHAPATCECTRESTLAML SLPRAWDYRWPHAANFCETREDEFSLAML SLPRAWDYRWPHAPARFCETRESLAML SLPRAWDYRWPHAPARFCETRESLAML SLPRAWDYRWPHAPARFCETRESLSTAFS*QAVCL SLPRAWDYRWPHAPATCETRESLATALFS*AQAVCL SLPRAWDYRWPHAPATCETCRESLAML SLPRAWDYRWPHAPATCETRESLATALFS*AQAVCL SLPRAWDYRWPHAANFCETRESLEGGESTAML SLPRAWDYRWPH	-							NIHLQGKLEDPAIRWQMALYKDLPNRTDDTS
RMAPLYNLPHARAVNI,FLQGYVERWIFTER YFEDKLIEDLARPGAEPPEDEGTKRYDPLH LILLFSRTALTEKCKLEEDFI, YMAY ADIMAK CHDEEDDDGEEVKSFEEKEMEKQKLLYQQ ARLHDRGAAEMVLQTISASKGETGPMVAAT LKLGIAILNGGNSTVQQKMLDYLRKKDVGI FQSLAGLMQSCSVLDLNAFERQNKAEGLGM VTEEGSGEKVLQDDEFTCDLFRIFQLLCEGH NSDFQNYLRTQTGNNTHISTVDYLLRVQ ESISDFYWYSGKDVIDEQGGRNFSKAIQVA KQVFNTLTEYIQGPCTGNQSLAHRRLWDAA VGFLHYFAHMQMKLSQDSSQIELLKELMDLA KDMYVMLLSMLEGNVVNGTIGKGMYDMLV ESSNNVEMILKFFDMFLKLKDLTSSDTFKEYI PDGKGVFFKRDFHKAMESHKHYTQSETFELL SCAETDENFLDYEEFKFHEPAKDIGFNV VLLTNLSEHMPNDTRLQTFLELAESVLNYFQ FLGRIEMGSAKRIERVYFEISESSTQWEKYG FLGRIEMGSAKRIERVYFEISESSTQWEKYG VRESKRQFFDVVNGEGEKEKMELFVNFGED TFEMQLAAQISESDLNERSANKESEKERFEI QGFRMAFFSILTVSSALFALRYNILTLMRMLS LKSLKQMKKVKKMTVEDMYTAFFSSYWS FMTLHIFVASVFRGFFRICSLLIGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEEGERKI LEAALPSEDLTDLKEITESDLLSDIFGLDLK EGGGYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQGKAKEEEKEEKEETKSEPEKAEGEDGE KEEKAKEDKGKKLRQUHTHRYGEPEVPESS FWKKIIAYQQKLLMYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVEGKELPTRSSSEAR KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRII PILHTVISFFCIIGYYCLKVPLVIFRREKEVARK LEFDGLITJTEOPSEDDLKGOMPRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRRISELIG MDKAALDFSDAREKKRKPKSDSLSAVINSID VKYQMWKLGVVYTDNSFLYLAUNTMSSVL GHYNNFFFFAAHLLDIAMGFKLRTILSSVTH NGKQLVLTVGLLAVVYYLTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIDPAGDSFLTRIIETDITFFFFIV VILLAIIQGLIIDAFGERDQQGQVEDMETIKC FICGIGNBYDTVPHGFETHTLQEHNLANYLF FLMYLRKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLIN 446 1796 A 3592 1 355 AGLELINSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPSSWDYRRWPPHAPAGLSLAML	- 1							OLVEUDODOVY ANYMINI BY ODYD ASSESSED
YFEDKLEDLAKPGAEPPEEDEGTKRYDPLH LILLFSRTALTEKCKLEDPLYMAYADIMAK CHDEEDDDGEEVKSFEEKEMEKQKLLYQQ ARLHDRGAAEMVLQTISASKGETGMVAAT LKLGIALINGGNSTVQKMLDYJKEKKDVGI FQSLAGLMQSCSVLDLNAFERQNKAEGLGM VTEEGSGEKVLQDDEFTLDLEDFRFLQLLCEGH NSDFQWYLRTQTIONNTTVMIISTVDYLLRVC ESISDFYWYYSGKDVIDEQGQNFISKAQVA KQVFNTLTEYIQGPCTGNQQSLAHSRLWDAN VGFLHVFAHMQMKLSQDSSQIELKELMDLL KDMVVALLSMLEGBVVNORTIGKQMVDML ESSNNVEMILKFFDMFLKLKDLTSSDTFKEFY PDOKGVFKROPHKAMESHKHVTQSETEFLI SCAETDENETLDYEEFVKRPHEPAKDIGFNV. VLLTNLSBHMPNDTRLQTFLELAESVLNYFQ FLGREIMGSAKRIERVFSSESSRTQWEKPC VKESKROFFDVVNEGGEKERMELFVNFCED TFENQLAAQISESDLNSANKEESKERPEI QGPRMAFFSILTVRSALFALRYNLTLMRMLS LKSLKKQMKKVKKMTVKDMVTAFSSYWS] FMTLHFVASVFRGFFRICSLLLGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEGGEKK LEAALPSEDLTDLKEL TESDLLSDIFGLDLKI EGGQYKLIPHNPNAGLSDLMSNPVPMEVOGE KFOGQKAKEEKEEKEETKSEPEKAEGEDGE KEEKAKEDKGKKRQLHTHRYGEPEVESS FWKKIIAYQQKLLNYFARNFYNMRNLALFV AFAINFILLFYKVSTSSVYEGKELPTRSSSENA KVTSLDSSSHRIJAVHYVLEESSGVMEFTVRIL PILHFVISFCIGTYYCLK VPL.VIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVWMKYGFFYGRINSELLG MDKAALDFSDAREKKKPKDSSLSAVINSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVJ, GHYNNFFFGAHLLDLAMGFKTLRTILSSVTH NGKQLVLTVGLLLAVVYJLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEPPAGDEYEVRIMEDITFFFFV VILLAIQGLIIDAFGELRDQQGQVKEDMETKC FICCIGIONDYDTYPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGGESYVWKMYQERCWE FFPAGDCTRKQYEVQHEMETKC FICCIGIONDYDTYPHGFTTHTLGHNLANYLF FLMYLINKDETEHTGGESYVWKMYQERCWE FFPAGDCTRKQYEVQHEMETKC FICCIGIONDYDTYPHGFTTHTLGHNLANYLF FLMYLINKDETEHTGGESYVWKMYQERCWE FFPAGDCTRKQYEVQHEMETKC FICCIGIONDYDTYPHGFTTHTLGHNLANYLF FLMYLINKDETEHTGGESYVWKMYQERCWE FFPAGDCTRKQYEVWEMETKC FICCIGIONDYDTYPHGFTTHTLGHNLANYLF FLMYLINKDETEHTGGESYVWKMYQERCWE FFPAGDCTRKQYEVWEMETKC SLPSSWOYNGWPHPHANGLIFCROESSLAML SLPSSWOYNGRWPHPHANGLIFCROESSLAML								
LILLIPSKTALTEKCKLEEDFI, WAYADIMAK, CHDEEDDDGEEWKSEEMEKOKLLYQQ ARLHDRGAAEMVLQTISASKGETGPMVAAT LKLGIAILNGGNSTVQQKMLDYLEKKDVOG FQSLAGLMQSCSVLDLNAFERQNKAEGLGM VTEEGSGEKVLQDDEFTCDLFRFLQLLCEGH NSDFQWYLRTQTGNNTWIISTVDYLLRVQ ESISDFYWYYSGKDVIDEQGGRNFSKAIQVA KQVPTILTEYIQGPCTOQSLAFRRLWDAA VGFLHVFAHMQMKLSQDSSQIELLKELMDLG KDMVVMLLSMLEGRVVNCTIGKGWMDMLV ESSNNVEMILKFFDMFLKLKDLTSSDTFKEYI PDGKGVFKRDFHKAMESHKHYTQSETFELL SCAETDENTLDYEEFVFHEPAKDIGFNV VLLTNLSEHMPNDTRLQTFLELAESVLNYPQ FLGREIMGSAKJREVYPEISESSRTQWEKPQ VKESKRQFIFDVVNEGGEKEKMELFVVFCED TFEMQLAAQISESDLNESANKESEKERPEL QGFRMAFFSILTVRSALFALRYNILTLMRMLS LKSLKKQMKKVKKMTVKDMVTAFFSSYWS FMTLHFVASVFRGFFRICSLLLGGSLVBGA KKIKVAELLANMPDFTQDEVRGDGEGGERK LEAALPSEDLTDLKELTEEDDLLSDIFGDLKS FGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE KTQGQKAKEEKKEEKEETKSEPEAGEDGE KEEKAKEDKGKKLRQUHTHRYGEPEVPESS FWKKIIAYQQKLLNYFARNFYNMRMLALFV AFANFILLFYKVSTSSVEGKELPTKSSSENA KVTSLDSSSHRIJAYHYVLEESSGYMEPTVRII PILHTVISFFCIIGYYCLKVPL-VIFREKEVARK LEFDGLYITEQFSEDDIKGQWDRLVNTQSFP NNYWDKFVKRKVMDKYGEFYGRDISELLG MDKAALDFSDAREKKKPKDSSLSAVINSID VKYQMWKLGVVFTDNSILYLAWYMTMSVJ- GHYNNFFFAAHLDLIAMGFTLRTILSSVTH NGKQLVLTVGLLAVVYYLYTVVAFNFFKKF YNKSEDGDTPDMKCDDMLTCYMHMYVGV RAGGGIGDEIEDPAGDEFTHTLGEINLANYLF FLMYLRIKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELINSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS-AQAVCL SLPRSWDYNRWPPHPANGCIFCRDESLAML								
CHDEEDDDGEEEVKSPEEKAMEKQKLLYQQ ARLHDRGAAEMVLOTISKGETGMVAAT LKLGIALINGGNSTVQQKMLDYLKEKKDVGI FQSLAGLMQSCSVLDLNAFERQNKAEGLGM VTEEGSGEVLQDDEFTCDLFRFLQLLCEGH NSDFQNYLRTQTGNTTTVNIISTTVYLLRVQ ESISDFYWYSGKDVIDEQGQRNFSKAIQVA KQVFNTLTEYIQGPCTGNQQSLAHSRLWDAN VGFLHVFAHMQMKLSQDSSQIELKELMDLU KDMVVALLSMLEGNVVNOTTIGKQMVDMLV ESSNNVEMILKFPDMFLKLKDLTSDFFKEYY PDOKGVFRKFDHKAMESHKHYTQSETEFIL SCAETDENETLDYEEFVKRFHEPAKDIGFNV, VLLTNLSHEMPNDTRLQTFLELASEVLNYFQ FLGREIMGSAKRIERVYFEISESSRTQWEKPQ VKESKROFFDVVNEGGEKERMELFVNFCED TFENQLAAQISESDLNSANKEESEKERPEI QGFRMAFFSILTVRSALFALRYNILTLMRMLS LKSLKXQMKZVKKMYVFEISESSRTQWEKPQ VKESKROFFDVVNEGGEKERMELFVNFCED TFENQLAAQISESDLNSANKEESEKERPEI QGFRMAFFSILTVRSALFALRYNILTLMRMLS LKSLKXQMKZVKKMYVFEISESSRTQWEKPQ VKESKROFFDVNFGGGEKERMELFVNFCED KESKROFFDVNEGGEKERMELFVNFCED KESKROFFDVNFGGGEKERMELFVNFCED KESKROFFDVNFGGGEKERMELFVNFCED KESKROFFDVNFGGGEKERMELFVNFCED VKESKROFFDVNFGGGEKERMELFVNFCED KESKROFFDVNFGGGEKERMELFVNFCED KESKROFFDVNFGGGEKERMELFVNFCED KESKROFFDVNFGGGEKERMELFVNFCED KESKROFFDVNFGGGEKERMELFVNFCES KKIKVAELLANMPDPTQDEVRGDGEGGERK LEAALPSDLTDLKELTESDLLSDIFGIDLKI EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQG KFOQKAKKEEKEKEKEKSPEKAGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKXIIAYQQKLLNYFARNFYNMRMLALFV AFANFELLFYKVSTSVYCKKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEFTVRIL PILHTVISFFCIGTYCICKSPEKAFEKEVARK LEFDGLYITTGPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFTGRDRISELIG MDKAALDFSDAREKKPKKDSSLSAVINSID VKYQMWKLGVVFTDNSFLYLLWYMTMSVL GHYNNFFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVYJLYTVVAFNFFRKF YNKSEDGDTPDMKCDMLTCYMFHMYVGV RAGGGIGDEIDPAGDEYEIYRIEDITFFFFV VILLAIQGLIIDAFGELRDQQEQVVEMMETICK FICCIGIONDYDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGGESYVWKMYQERCWE FFPAGDCERKQYEDQLS 446 1796 A 3592 1 355 AGLELINSDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPSSWOYNRWPPHPANFCIFCRDESSLAML	-							
ARLHDRGAAEMVLQTISASKGETGPMVAAT LKLGIALINGGNSTVQKMLDYLKEKKDVGI PQSLAGLMQSCSVLDLNAFBRQNKAEGLGM VTEEGSGEKVLQDDEFTCDLLCEGH NSDFQNYLRTQTGNNTTVNIIISTVDYLLRVQ ESISDFYWYYSGKDVIDEQQGRAFISKAIQVA KQVFNTLTEYIQOPCTORQQSLAHSRLWDAN VOFLHVFAHMQMKLSQDSSQIELLKELMDLA KDMVVMLLSMLEGNVVNGTIGKQMVDMLV ESSNNYEMILKFFPMFLKLKDLTSSDTYKEYI PDGGGVIFKRDFHKAMESHKHYTQSETEFLL SCAETDENETLDYEEFVKRFHEPAKDIGFNV VLLTNLSEHMPNDTTRLQTFLELAESVLNYFQ FLGRIEIMGSAKRIERYYFEISESSRTQWEKPQ VKESKROFIFDVNTGGGEKEKMELFVHFOED TIFEMQLAQISESDLNERSANKEESEKERPEI QGPRMAFFSILTVRSALFALRYNILTLMRMLS LKSLKKQMKKVKKMTVKDMVTAFFSSYWSI FMTLLHFVASVFRGFFRIICSLLGGSLVEGA KKIKVABLLANMPDPTQDEVRGDGEGEGRKI LEAALPSEDLTDLKELTEESDLLSDIFGDLKS FGGQYKLIPHNPAGLSDLMSNPVPMPEVQE KFQEQKAKEEKEEKEETKSEPEKAEGEDGE KEEKAKEDKGKQKLRQUHTHRYGFPEVPESA FFKKILAYQQKLLNYFARNFYNMRMLALFV AFANFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRI PILHTVISFFCIIGY YCLKVPLVIFRREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFYKRKVMDKYGFFYGRDRISELG MDKAALDFSDAREKKKPKCDSSLSAVLNSID VKYQMWKLGVVTDNSFLYLAWYMTMSVL GHYNNFFFAAHLDIAMGFKTLRTILSSVTH NGKQLVLTVGLIAAVVVYLYTVAFNFRKF YNSSEDGJTPDMKCDDMLTCYMFHRWYGV RAGGGIGDEEDPAGDEYEIYRIFDITFFFVI VILLAIIQCILIDAFGELRIOQEGVKEDMETT, NGKQLVLTVGLIAAVVVYLYTVAFNFRKF YNSSEDGJTPDMKCDDMLTCYMFHRWYGV RAGGGIGDEEDPAGDEYEIYRIFDITFFFFVI VILLAIIQCILIDAFGELRIOQEGVKEDMETT, NGKQLVLTVGLIAAVVVYLYTVAFNFRKF YNSSEDGJTPDMKCDDMLTCYMFHRWYGV RAGGGIGDEEDPAGDEYEIYRIFDITFFFFVI VILLAIIQCILIDAFGELRIOQEGVKEDMETT, SSTHMYGPFGFTHTLQEHNLANTYLF FLMYLINKDETEHTGGESYVWKMYQERCWE FFPAGDCFRKAYEDQLN 446 1796 A 3592 1 355 AGLELLNSDDPPALASGSAGITGVTRTPSLFFF DTVLLCCSGWSAVAPSRLTAALFS*QAACL SPRSWVYRWPPPHPANFCIFCRDEFSLAML	ſ							
LKLGIAILNGGNSTVQQKMLDYTLKEKKDVGI FQSLAGLMQSCSVLDLNAFERQNKAEGLGM VTEEGSGEKVLQDDETCDLFRILQLLCEGH NSDFQNYLRTQTGNNTTVNIIISTVDYLLRVQ ESISDFYWYYSGKDVIDEQQQRNFSSKAIQVA KQVFNTLTEYIQQFCTGNQQSLAHSRLWDAN VGFLHVFAHMQMKLSQDSSQIELKELMDLL KDMVVMLLSMLEGNVVNGTIGKQMVDMLV ESSNNVEMILKFFDMFLKLKDLTSSDTFKEYY PDGKGVFKRDFHKABHKHYTQSTETFLL SCAETDENETLDYEEFVKRFHEPAKDIGFNV, VLLTNLSEHMPNDTRLQTFLELAESVLNYFQ FLGRIEMGSAKRIERYYFEISESSRTQWEKPQ FLGRIEMGSAKRIERYYFEISESSRTQWEKPQ VKESKRQFIFDVVNEGGEKEKMELFVNFCED TIFEMQLAAQISESDLNERSANKEESEKERPEI QGPRMAFFSILTVSASLFALRYNILTLMRMLS LKSLKKQMKKVKKMTVKDMVTAFFSSYWS; FMTLLHPVASVFRGFTIGSLLLGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEEGERKI LEAALPSEDLTDLKELTEESDLADIFGLDLS FMCLIHPNVANGLSDLMSNPVPMPEVQE KFQEQKAKEEEKEEKEETKSEPPKARGEDOE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKXIIAYQQKLLNYFARNIFYNNMMLALFV AFANFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVTYLEESSGYMEPTVRII PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFGGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRXWMDKYGEFYGRDRISELLG MDKAALDFSDAREKKPKKDSLSAVLNSID VKYQMWKLGVVTDSFLYLAWYMIMSVL, GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFFNFRKF YNSEDGOTPDMKCDDMLTCYMFHRWYGV RAGGGIGDEIEDPAGDEYSIYRIEDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEBNLANYLF FLMYLINKDETEHTGQESYWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELINSDDPPALASQSAGITGVTRTPSLFF DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PIPPANFCIFCRDEGSLAML	ì							
FQSLAGLMQSCSVLDLNAFERQNKAEGLGM VTESGGEKVLQDDETTCDLFRFLQLLCEGH NSDFQNYLRTQTGNNTTVNIISTVDYLLRVQ ESISDFYWYYSGKDVIDEQQGRAFSKAIQVA KQVFNTITEYIQGPCTQQSLAHSRLWDAN VGFLHVFAHMQMKLSQDSSQIELLKELMDLL KDMVVMLLSMLEGNVVNGTIGKQMVDMLL ESSNNVEMILKFFDMFLKLKDLTSSDTFKEYI PDGKGVIFKRDFHKAMESHKHYTQSETEFLL SCAETDENETLDVEEVKRPHEPAKDIGFNV, VLLTNLSEHMPNDTRLQTFLELAESVLNYFQ FLGRIEMGSAKRIERVYFEISESSRTQWEKPQ VKESKRQFIEDVVNEGGEKEKMELFVNFCED TIFEMQLAAQISESDLNERSANKEESEKERPEJ QGPRMAFFSILTVSAFLAENVILITLMRMLS LKSLKKQMKKVKKMTVKDMVTAFFSSYWSI FMTLLHFVASVFRGFFRIICSLLGGSLUEGG KEKARLLANMPDPTQDEVRGDGEGEGRK LEAALPSEDLTDLKELTEESDLADIFGLDLKS EGGQVKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEQKAKEEKEKEEKEKEFKSPEKAEGEDGE KEEKAREDKGKQKLRQHTHRYGFEPEVPESS FWKKILAYQQKLLNYFARNFYNMRMLALFV AFANNFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRILAVYLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLKVFLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQMDRLVINTQGFP NNYMDKFVKRKVMVGGFYGORDRISELLG MDKAALDFSDAREKKKPKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVAFFRKFK YNSEDGDTPDMKCDDMLTCYMFHRWYGV RAGGGIGDEEDPAGDEVEIYRIEDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETTY VILLAIIQGLIIDAFGELRDQQEQVKEDMETTY VILLAIIQGLIIDAFGELRDQQEQVKEDMETTY FILMYLINKDETEHTGQESVWKMYQERCWE FPFAGDCFRKQVEDQLN 446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*QAVCL SLPRSWDYRFWRPHAPNFICFCREDESLAML	1							
VTEEGSGEKVLQDDEFTCDLFRFLQLLCEGH NSDFQNYLRTQTGNNTTVNIISTVDYLLRVQ ESISDFYWYYSGKDVIDEQGQRNFSKAIQVA KQVFNTLTEYIQGPCTGNQSLAHSRLWDAN VGFLHVFAHMQMKLSQDSSQIELIKELMDLL KDMVVMLLSMLEGNVVNGTIGKQMVDMLV ESSNNVEMILKFFDMFLKLKDLTSSDTFKEYI PDGKGVIFKRDFHKAMESHKHYTQSETFFLL SCAETDENETLDYEEFVKRFHEPAKDIGFNV. VLLTHLSHHMPNTRLQTFLELAESVLNYFQ FLGRIEIMGSAKRIERVYFEISESSRTQWEKYG VKESKRÖFFDVVNEGGEKEKMELFVNFCED TIFEMQLAAQISESDLNERSANKESSKERPEL QOPRMAFFSILTVRSALFALRYNILTLMRMLS LKSLKKQMKVKKKNTVKDMVTAFFSSYWSI FMTLLHFVASVPRGFFRICSLLLGGSLVEGA KKIKVAELLANMPDPTODEVRGDGEGGRKI LEAALPSEDLTDLKELTESDLLSDIFGLDLKI EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEQKAKEEKKEEKEETKSPFEKAEGEDGE KEEKAKEDKGKQLRQHTHRYGFEPVPESS, FWKKIIAYQQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSYMEPTVRIL PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFOBL YITEOPSEDIKGOWDRLVINTGSFP NNYWDKFVKRVMDKYGEFYGRDRISELIG MDKAALDFSDAREKKRPKKDSSLSAVLNSID VKYQMWKLGVVYLVYTVVAFNFFRKF YNKSEDGDTFDMKCDDMLTCTMFHMYGGV RAGGGIGDEIDPAGDEFURTIESTTFFV VILLAIQGLIDAFGERDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESVYWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELNSDDPPALASQSAGITGVTRTPSLFFF DTVLLCCSGWSAVAPSRITAALFS*AQAVCL SLPRSWDYRKWPYPHPANFCICCRDEFSLAML	.							
ESISDFYWYSGKDVIDEQGQRNFSKAIQVA KQVFNTLTEYIQGPCTGNQQSLAHSRILWDAN VGFLHVFAHMQMKLSQDSSQIELLKELMDLL KDMVVMLLSMLEGNVVNGTIGKQMVDMLV ESSNNVEMILKFDFMFLKLKULTSSDTFKEYT PDGKGVIFKRDFHKAMESHKHYTQSETEFLL SCAETDENETLDYEEFVKRFHEPAKDIGFNV. VLLTINLSEHMPNDTRLQTLELAESVLNYRG FLGRIEMGSAKRIERVYFEISESSRTQWEKPG VKESKROFIFDVNTEGGEKKEMELFVNFGED TIFEMQLAAQISESDLNERSANKESSEKERPEL QGPRMAFFSILTVRSALFALRYNILTLMRMLS LKSIKKQMKKVKMTVKDMVTAFFSSYWSI FMTLLHFVASVFRGFFRIICSLLLGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEGGERKI LEAALASPEDLTIDLKELTESDLLSDIFGLDLKI EGGQYKLIPHPNAGLSDLMSNPVPMPEVQE KFQEQKAKEEKEKEEFKEFKSPEKAEGEDGE KEEKAKEDKGKOKLRQLHTHRYGDEFVPESS FWKKIIAYQQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRII PILHTVISFFCIIGYYCLKVPLVIFKREKEVARN LEFDGLYTTEQPSEDDIKGQWDRLVINTGSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDEDDIKGGWDRLVINTGSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKKRKKDSSLSAVLNSID VKYQMWKLGVVTNSFLYLAWYMTMSVL, GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYYGV RAGGGIGDEIEDPAGDEYBIYRIIEDITTFFFFVI VILLALIQGLIDDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANVLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFFRKQYEDQIN 446 1796 A 3592 1 355 AGLELINSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRITAALFS*AQAVCL SLPRSWDYRKWPYPHANFCIFCRDEFSLAML	1							VTEEGSGEKVLQDDEFTCDLFRFLQLLCEGH
KQVFNTLTEYIQQFCTGNQQSLAHSRLWDAY VGFLHVFAHMQMKLSQDSSQIBLLKELMDLG KDMVVMLLSMLEGNVVNGTIGKQMVDMLV ESSNNVEMILKFFDMFLKLKDLTSSDTFKEYI PDGKGVIFKRDFHKAMESHKHYTQSETEFLI SCAETDENETLDVEEFVKRFHEPAKDIGFNV. VLITNLSEHMPNDTRLQTFLELAESVLNYFQ FLGREIMGSAKRIERVYFEISESSRTQWESSSSRTOWESSEKERPEI QOPRMAFFSLTVRSALFALRYNLTLMRMLS LKSLKKQMKKVKKMTVKDMVTAFFSSYWSI FMTLLHFVASVFRGFFRICSLLLGGSLVEGA KKIKVAELLANMPDPTQDEVFRGDEEGGERKI LEAALPSEDLTDLKELTEFSDLLSDIFGLDLKI EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQCKAKEEKEEKEETKSFPEKAEGEDGE KFEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKIIAYQQKLLNYFARNFYNMRMLALFV AFANIFILLFYKVSTSSVVEGKELFIRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRII PILHTVISFFCIIGYVCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRVMDKKYGEFYGRDRISELIG MDKAALDFSDAKKKPKKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLAVVYVLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGBEEDPAGDEYEIYRIIFDITFFFVI VILLAIQGLIIDJAFGELRDQGQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGGESYVWKMYQERCWE FFPAGDCFRKQVEDQLN 446 1796 A 3592 1 355 AGLELINSDDPPALASQSAGITGVTRTPSLFF DTVLLCCSGWSNAVAPSRLTAALFS*AQAVCL SLPRSWDYRRWPPHPANFCIFCRDESLAML	١							NSDFQNYLRTQTGNNTTVNIIISTVDYLLRVQ
VGFLHVPAHMQMKLS,ODSSQIELLKELMDLI KDMVVMLLSMLEGNVVNGTIGKQMVDMLV, ESSNNVEMILKFFDMFLKLKDLTSSDTFKEYI PDGKGVIFKRDFHKAMESHKHYTQSETEFLL SCAETDENETLDYEEFVKRFHEPAKDIGFNV, VLLTNLSEHMPNDTRLQTTLELAESVLNYFQ FLGRIEIMGSAKRIERVYTEISESSRTQWERGG VKESKROFIFDVVNEGGEKEKMELFVNFCED TIFEMQLAAQISESDLNERSANKEESEKERPEI QGPRMAFFSLTVRSALFALRYNLTTLMRMLS, LKSLKKQMKK VKKMTVKDMVTAFFSSYWSI FMTLLHFVASVFRGFFRIICSLLLGGSLVEGA KKIKVAELLANMPDPTODEVRGDGEEGERKI LEAALPSEDLTDLKELTEESDLLSDIFGLDLKI EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEQKAKEEEKEEEFKSEPEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKIIAYQQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELFITSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL PILHTVISFFGIGYYCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKEPKKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYYGV RAGGGIGEIEDPAGDEYEIYRIIFDITFFFVI VILLAIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTTQESYVWKMYQERCWE FFPAGDCFRKQVEDQLN 446 1796 A 3592 1 355 AGLELINSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYKRWPPHPANFCIFCRDE/SLAML	1							ESISDFYWYYSGKDVIDEQGQRNFSKAIQVA
KDMVVMLLSM.EGRVVNGTIGKQMVDMLV ESSNNVEMILKFFDMFLKLKDLTSSDTFKEYI PDOKGVIFKROPHKAMESHKHYTQSETEFILL SCAETDENETLDYEEFVKRFHEPAKDIGFNV, VLLTNLSEHMPNDTRLQTFLELAESVLNYFQ FLGRIEIMGSAKRIERVYFEISESSRIQWEKPQ VKESKROPIFDVVNEGGEKEKMELFVNFCDE TIFEMQLAAQISESDLNERSANKEESEKERPEI QGPRMAFFSILTVRSALFALRYNILTLMRMLS LKSLKKQMKKVKKMITVKDMVTAFFSSYWS FMTLLHFVASVFRGFFRICSLLIGGSLVEGA KKIKVAELLANMPDPTQDEVRODGEEGERKI LEAALPSEDLTDLKELTEESDLLSDIFGLDLKI EGGGYKLIPHPNAGLSDLMSNPVPMPEVQE KFQEQKAKEEKEEKEETKSEPEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESS FWKKIIJAYQKLINYPARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRII PILHTVISFFGUYCLK VPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELIG MDKAALDFSAREKKFKKGSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLTLSISSTH NOKQLVLTVGLLAVVVYLTVVAVAFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIDPAGDEYEIYRIFDITFFFFV VILLAIIQGLIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGGELLNSDDPPALASQSAGITGVTRTPSLFFP DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRWPPHPANFCIFCRDESSLAML	ı							KQVFNTLTEYIQGPCTGNQQSLAHSRLWDAV
ESSNNVEMILKFFDMELKLKDLTSSDTFKEYT PDGKGVIFKRDFHKAMESHKHYTQSETEFLL SCAETDENETLDYEEFVKRFHEPAKDIGFNV. VLLTNILSEHMPNDTRLQTFLELAESVLNYFQ. FLGRIEMGSAKRIERVYFEISESSRTQWEKPQ VKESKRQFIITDVVNEGGEKEKMELFVNFCED TIFEMQLAAQISESDLNERSANKEESEKERPEJ QGPRMAFFSILTVRSALFALRYNILTLMRMLS LKSLKQMKKVKKMITVKDMYTAFFSYWS FMTLLHFVASVFRGFFRIICSLLLGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEEGERK LEAALPSEDLTDLKELTESPDLLSDIFGJLKI EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE KTQEQKAKEEKEEETSSEPEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKXIIAYQQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTISSSSMA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRII PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKKPKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVYYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIDPAGDEYSIRIFDTFFFFVI VILLAIIQGLIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEINLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCTRRQYEDQLN 446 1796 A 3592 1 355 AGLELNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRWPPHPANFCIFCRDE/SLAML	1							
PDGKGVIFKRDFHKAMESHKHYTQSETEFLL SCAETDENETLDYEEFVKRFHEPAKDIGFNV, VLLTNLSEHMPNDTRLQTFLELAESVLNYFQ FLGRIEIMGSAKRIERVYFEISESSRTQWEKPQ VKESKRÇIFIDVVNEGGEKEKMELFVNFCED TIFEMQLAAQISESDLNERSANKEESEKERPEI QGPRMAFFSILTVRSALFALRYNILTLMRMLS LKSLKKQMKKVKKMTVKDMVTAFFSSYWS FMTLLHFVASVFRGFFRICSLLLGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEEGERKI LEAALPSEDLTDLRELTEESDLLSDIFGLDLKI EGGYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEQKAKEEKEEKEETKSEPEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKIIAYQQKLLNYFARNFYNMRMLALFV AFANFILLFYKVSTSSVVEGKELPTISSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRII PILHTVISFFCIIGYYCLKVPLLVIFKREKEVAAK LEFDGLYITTEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKPKKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPBMCDDMLTCYMFHNYVGV RAGGGIGDEEDPAGDEVEIVRIIFDITFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDIVPHGFETHTLQEINLANYLF FLMYLNIKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRWPPHPANFCIFCRDESSLAML	1							` · · · · · · · · · · · · · · · · · · ·
SCAETDENETLDYEEFVKRPHEPAKDIGFNV, VLLTNLSEHMPNDTRLQTFLELAESVLNYFQ FLGRIEIMGSAKRIERVYFEISESSRTQWEKRY VKESKRQFIFDVVNEGGEKEKMELFVNFCED TIFEMQLAAQISESDLNERSANKEESEKERPEI QGPRMAFSIETUVRSALFALRYNILTLMRMLS LKSLKKQMKKVKKMTVKDMVTAFFSSYWSI FMTLLHFVASVFRGFFRICSLLIGGSLVEGA KKIKVAELANMPDPTQDEVRGDGEEGERKI LEAALPSEDLTDLKELTESDLLSDIFGLDLKI EGGQVKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEQKAKEEEKEEFIKSEPEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESS FWKKIIAYQQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRII PILHTVISFFCIIGYYCLKYPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKSVKRKVMMKYGEFYGRDRISELIG MDKAALDFSDAREKKRPKKDSLSAVLNSID VKYQMWKLGVVTDNSFLYLAWYMTMSVL GHYNNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYLFVIIEDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYPDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTIGQESYVWKMYQERCWE FFPAGDCFRKQVEDQLN 446 1796 A 3592 1 355 AGGELLNSDDPPALASQSAGITGYTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRWJPPHPANFCIFCRDESLAML	1							
VLLTNLSEHMPNDTRLQTFLELAESVLNYFQ FLGRIEIMGSAKRIERVYFEISESSRTQWEKPG VKESKRQFIFDVVNEGGEREKMELFVNFCED TIFEMQLAAQISESDLNERSANKEESEKERPEI QGPRMAFFSILTVRSALFALRYNILTLMRMLS LKSLKKQMKKVKKMTVKDMVTAFFSSYWSI FMTLLHFVASVFRGFFRICSLLLGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEEGERKI LEAALPSEDLTDLKELTEESDLLSDIFGLDLKI EGGQVKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEQKAKEEEKEEKEETKSEPEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKIIAVQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRII PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKRPKDSSLSAVLNSID VKYQMWLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVYYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIEDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANTLF FICMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGGELINSDDPPALASQSAGITGYTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRWJPPHPANFCIFCRDESLAML	J							
FLGRIEIMGSAKRIERVYFEISESSRTQWEKPÖ VKESKRQFIFIDVVNEGGEKEKMELFVNFCED TIFEMQLAAQISESDLNERSANKEESEKERPEI QGPRMAFFSILTVRSALFALRYNILTLMRMLS LKSLKKQMKKVKKMTVKDMVTAFFSSYWS; FMTLLHFVASVFRGFFRIICSLLLGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEGERKI LEAALPSEDLTDLKELTEESDLLSDIFGLDLKI EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEQKAKEEKEEKEETKSEPEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKIIAYQQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLK VPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRVMDKYGEFYGRDISELLG MDKAALDFSDAREKKPKKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRWPPHPANFCIFCRDE/SLAML	1							· · · · · · · · · · · · · · · · · ·
VKESKROFIFDVVNEGGEKEKMELFVNFCED TIFEMQLAAQISESDLNERSANKEESEKERPER QOPRMAFFSILTVRSALFALRYNILTLMRMILS LKSLKKQMKKVKMTVKDMVTAFFSSYWS) FMTLLHFVASVPRGFFRIICSLLLGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEEGERKI LEAALPSEDLTDLKELTESDLLSDIFGLDLKI EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEQKAKEEEKEEKEETKSEPEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKXILAYQQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRII PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKPKKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELINSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SJEPRSWDYRRWPPHPANFCIFCRDE/SLAML	- [,	
TIFEMQLAAQISESDLNERSANKEESEKERPEI QGPRMAFFSILTVRSAALFALRYNILTLMRIMLS LKSKKQMKKVKKMTVKDMVTAFFSSYWS FMTLLHFVASVFRGFFRIICSLLLGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEEGERKI LEAALPSEDLTDLKELTESDLLSDIFGLDLKI EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEQKAKEEKEEKEETKSEPEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKIIAYQQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRII PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKPKKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEVEIYRIIFDITFFFFVI VILLAIIQGLIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPDFFFTHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELINSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRWIPPHPANFCIFCRDESSLA/ML	1			·				
QGPRMAFFSILTVRSALFALRYNILTLMRMI.S LKSLKKQMKKVKKMTVKDMVTAFFSSYWSI FMTLLHFVASVFRGFFRIICSLLLGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEEGERRIX LEAALPSEDLTDLKELTEESDLLSDIFGLDLKI EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEQKAKEEKEKEETKSEPEKAEGEDGE KEEKAKEDKGKQKLRQLHTTHRYGEPEVPESA FWKKIIAYQQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKPKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELINSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRWPPHPANFCIFCRDE/SLAML	7						1	TIFEMOLAAOISESDI.NERSANKEESEKERPEE
LKSLKKQMKKVKMTVKDMVTAFFSSYWSJ FMTLLHFVASVFRGFFRIICSLLLGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEEGERKJ LEAALPSEDLTDLKELTEESDLLSDIFGLDLKEL EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEQKAKEEEKEEKEFTKSEPEKAEGEIGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKIIAYQQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRII PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKKPKKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEEDPAGDEYEIYRIIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELINSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRWPPHPANFCIFCRDE/SLA/ML	1	}		[İ	
FMTLLHFVASVFRGFFRIICSLLLGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEEGERRI LEAALPSEDLTDLKELTEESDLLSDIFGLDLKI EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEKAKEEKEEKEEKEETKSEPEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKIIAYQQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRII PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKKPKKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML	1	ļ		ļ		J		
KKIKVAELLANMPDPTQDEVRGDGEEGERKI LEAALPSEDLTDLKELTEESDLLSDIFGLDLKI EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEQKAKEEEKEEKETKSEPEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKIIAYQQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRII PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKKPKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGIGDEIEDPAGDEYEIYRIIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRWPPHPANFCIFCRDE/SLA/ML								
EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEQKAKEEKEEKEETKSEPEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKIIAYQQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRII PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKKPKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKO FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML	1							KKIKVAELLANMPDPTQDEVRGDGEEGERKP
KFQEQKAKEEEKEETKSEPEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKXIIAYQQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKKPKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKO FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRWPPHPANFCIFCRDE/SLAML	1	ļ				ľ		LEAALPSEDLTDLKELTEESDLLSDIFGLDLKR
KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKIIAYQQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRII PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFF NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKPKKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML	1	Ì		- 1				EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE
FWKKIIAYQQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRII PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFF NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKKPKKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML	1		}	1		İ	.]	KFQEQKAKEEEKEEKEETKSEPEKAEGEDGE
AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRII PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKKPKKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKO FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML				j			,	KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA
KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRII PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKKPKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML		}		!				FWKKIIAYQQKLLNYFARNFYNMRMLALFV
PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKKPKDSSSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML			[ĺ	[İ		
LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKKPKKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELINSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML				· [ĺ	{		
NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKKPKKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELINSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML				Ì			ļ	
MDKAALDFSDAREKKKPKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKO FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML		ł	1	ŀ	1	1		
VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHY\NNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKO FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML	1	Į		j	})		
GHYNNFFRAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKO FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML]]		ì		1	
NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML	Ţ		j					
YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML	1	1	- 1	1	Ì	1		
RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKO FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML		i	1	ľ	1	1	· 1	
VILLAIIQGLIIDAFGELRDQQEQVKEDMETKO FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML		1	- 1	1	Ì	1	l	
FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML			ł	j	ļ		ĺ	VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC
FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML		j	ſ	ļ	j	j	ļ	FICGIGNDYFDTVPHGFETHTLQEHNLANYLF
446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML		J	1	j		}]	FLMYLINKDETEHTGQESYVWKMYQERCWE
DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML	L							
SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML		446	1796	A	3592	1	355	AGLELLNSDDPPALASQSAGITGVTRTPSLFF*
	1	}	}	Į.	Ì	l	ł	
1 DOLLAR MODERN AND AND AND AND AND AND AND AND AND AN		-]	1		1	ļ	
L	L							PRLVSNSWTQAILLPRPPKMLGLQV

SEO ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide .	D=Aspartic Acid. E=Glutamic Acid.
nucl-	peptide	nou	in in	nucleotide	location	1
eotide	seq-	1	USSN	1	1	F=Phenylalanine, G=Glycine, H=Histidine,
1 ' '	uence		09/496	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	i	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	ļ		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
Į.		l		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
ì	Ì	1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
ł	l	ì	1	peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
447	1797	A	3598	1202	1070	LFVGGGPICPEGASGFAPGPAPAPRVGVDAEV
ľ			1	i		GR*V*GAAASQGA/GSLRPRPTGPGHPGAWL
	!					QVWGAAAVCAGPAM*/AVRAKRGPRAG*EP
1	1	ł		ł	1	NSPWRSGVLAA\RAVGAGPWP*P*PGCS*ARG
ł	l	1	l	ľ		PSSRSAPGLASGPAAPLLQGVHSSAGPLLCYI
L		İ		ļ		NGTLALGLKP**AWGWGEWRPKG
448	1798	Α	3604	3115	557	FRRKGGGGPKDFGAGLKYNSRHEKVNGLEE
						GVEFLPVNNVKKVEKHGPGRWVVLAAVLIG
1	ŀ]		LLLVLLGIGFLVWHLQYRDVRVQKVFNGYM
1	1	l		}	ì	RITNENFVDAYENSNSTEFVSLASKVKDALKL
1 .	1	{	}	}	ļ	LYSGVPFLGPYHKESAVTAFSEGSVIAYYWSE
	}	İ				FSIPOHLVEEAERVMAEERVVMLPPRARSLKS
		ļ				FVVTSVVAFPTDSKTVQRTQDNSCSFGLHAR
ĺ	1	ĺ			1	GVELMRFTTPGFPDSPYPAHARCQWALRGD
1						ADSVLSLTFRSFDLASCDERGRHLV\TVYNT\L
1	Į.	ł				SPMEPHA\LVQLCGTYPPSYNLTFHS\S\QNVL
}	Į	ļ]			LITLITNTERRHPG\FEATFFQLPRMSSCGGRL
Į	İ					RKAQGTFNSPYYPGHYPPNIDCTWNIEVPNN
	{	ľ		ļ		QHVKVRFKFFYLLEPGVPAGTCPKDYVEING
1	ĺ					EKYCGERSQFVVTSNSNKITVRFHSDQSYTDT
1	ļ					GFLAEYLSYDSSDPCPGQFTCRTGRCIRKELR
						CDGWADCTDHSDELNCSCDAGHQFTCKNKF
					j	CKPLFWVCDSLNDCGDNSDEQGCSCP\AQTF
						RCSNGKCLSKSQQCNGKDDCGDGSDEASCP
						KVNVVTCTKHTYRCLNGLCLSKGNPECDGK
						EDCSDGSDEKDCDCGLRSFTRQARVVGGTD
ļ .						ADEGEWPWQVSLHALGQGHICGASLISPNWL
						VSAAHCYIDDRGFRYSDPTQWTAFLGLHDQS
				:		QRSAPGVQERRLKRIISHPFFNDFTFDYDIALL
						ELEKPAEYSSMVRPICLPDASHVFPAGKAIWV
}						TGWGHTQYGGTGALILQKGEIRVINQTTCEN
						LLPQQITPRMMCVGFLSGGVDSCQGDSGGPL
l i			,			SSVEADGRIFQAGVVSWGDGCAQRNKPGVY
						TRLPLFRDWIKENTGV
449	1799	A	3618	2	613	FVSGSPWRMDGSTERLEARRPAGRLPWSSRQ
1						EMTRRPSLMAGRQHGWSAQQSATVANPVPG
l i				i		ANPDLLPHFLGEPEDVYIVKNKPVLLVCKAV
i 1						PATQIFFKCNGEWVRQVDHVIERSTDGSSGLP
j l			. 1			TMEVRINVSRQQVEKVFGLEEYWCQCVAWS
	ĺ	ĺ		ĺ	ļ	SSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSL
[[ļ		ł	Ì	l	EQGIVLPCRPPEGIPPAE
450	1800	A	3620	1	2676	MEPSLGQGMDLTCPFGVSPACGAQASWSIFG
}		1	- 1		i	ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE
, ,				ĺ		PQAAQSPAGQQGPPTAGVSCSPTPTIVLTGDA
1				ļ	Į.	TSPEGETDKNLANRVHSPHKRLSHRHLKVST
			ĺ			ASLTSVDPAGHIDLVNDQLPDISISEEDKKKN
[í	ł	1	ľ		LALLEEAKLVSERFLTRRGRKSRSSPGDSPSA
	ļ					VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS
	ļ	.	ŀ			SPHPGEPNVPKGLADRKONDORKVSOGRLAP
]			:			RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA
[[- 1	j	İ		ľ	PVTNSSGKMALNSPQPGPVESELGKQLLKTG
		1			ł	WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP
		ļ	ł			MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE
)	}	J				PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK
		ì		1	İ	APGLKDFQIQVQPVRMQKLTKLREEHILMRN
ļ İ		ł		İ		QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE
	ł	ł	- 1	Ì		ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP
	ŀ	ł	}	1	1	LTEKEVENVFVQLSSAFRNDSYTLESRINQAE
	1	ļ	ĺ			RERNLTEENTEKELENFKASITSSASLWHHCE
Ll						HRETYQKLLEDIAVLHRLAARLSSRAEVVGA

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion VRQEKRMSKATEVMMQYVENLKRTYEKDH AELMEFKKLANQNSSRSCGPSEDGVLRTARS MSLTLGKNMPRRVSVAVVPKFNALNLPGQ TPSSSSIPSLPALSESPNGKGSLPVTSALPALLE NGKTNGDPDCEASAPALTLSCLEELSQETKA RMEEEAYSKGFQEGLKKTKELQDLKEEEEEQ KSESPEEPEEVEETEEEEKDPRSSKLEELVHFL QVMYPKLCQHWQVIWMMAAVMLVLTVVL
L					[GLYNSYNSCAEQADGPLGRSTCSAAQKDSW WSSGLQHEQPTEQ
451	1801	A	3623	504	198	QLIQHQTVHTGRKLYECKECGKAFNQGSTLI RHQRIHTGEKPYECKVCGKAFRVSSQLKQHQ RIHTGERPYQCKELKGRGAEMLAVLAVKEQ NRTPVNYGK
452	1802	A	3628	2	195	MTCLHSAKAFHY*SSCSFSCEEGFALIGPEVV QCTALGVWTAPAPVCIAVQCQHLEALNEGT MG*DYPFTAFAYGSSCKYECHTVYRVRGLD MLHSRGCYLWNGHFTT*EAISCEPLERPCH*S V*CSFSCEEGFALIGPEVVQCTALGVWTAPAP VCIAVQCQHLEALNEGTMG
453	1803	A	3637	662	142	IQAKGLGIWHVPNKSPMQHWR\KGSLLRYRT DTGFLQTLGHNLLGIYQKYPVKYGEGKCWT DNGPVIPVVYDFGDAQKTASYYSPYGQREFT AGFVQFRVFNNERAANALCAGMRVTGCNTE HHCIGGGGYFPEASPQQCGDFSGFDWSGYGT \\(\text{VFQYSSSREITE\AAVLLFYR}\)
454	1804	A	3641	1	362	TQVHPAMLGLDELGRSGCGHCTQADLRFGD AAGRDPGQDNDRNTAEPAFPPPRVMAAAA ALRAPAQSSVTFEDVAVNFSLEEWSLLNEAQ GCLYHDVMLETLTLISSLGKVLILNCDLS
455	1805	Ā	3646	2	414 .	AAAGRGASGALTGEGGGEQGRRVGLGSRAH
						SLLLGPTFNSCQVSSQPPRVAGLGLPLKHEPS RPQPPSPRGPRTVRAGVPGAHPQDTPCPEFVR PRKVPLVGEAPGLPPEERSRGWRRDTPGLQE SRVRAPSYDDIT
456	1806	A	3656	396	8	QIVSFNSYLTLYTKNNLKSMKDLNVNTEMIK LLELKNIHNLG*AKFFLN*IQKALIKRKILIHW P/LIKIK/SFCSLSDTIKKMKRQTIVWEQTFIIHI SVKELVSRIYEAFLQFNKTVNRPVFDIKKEQK F
457			3660		1961	SEAKLGGPTGMDLWQLLLTLALAGSSDAFSG SEATAAILSRAPWSLQSVNPGLKTNSSKEPKF TKCRSPERETFSCHWTDEVHHGTKNLGPIQLF YTRRNTQEWTQEWKECPDYVSAGENSCYFN SSFTSIWIPYCIKLTSNGGTVDEKCFSVDEIVQ PDPPIALNWTLLNVSLTGIHADIQVRWEAPRN ADIQKGWMVLEYELQYKEVNETKWKMMDP ILTTSVPVYSLKVDKEYEVRVRSKQRNSGNY GEFSEVLYVTLPQMSQFTCEEDFYFPWLLIIIF GIFGLTVMLFVFLFSKQQRIKMLILPPVPVPKI KGIDPDLLKEGKLEEVNTILAIHDSYKPEFHS DDSWVEFIELDIDEPDEKTEESDTDRLLSSDH EKLHINLGVKDGDSGRTSCCEPDILETDFNAH DIHEGTSEVAQPQRLKGEADLLCLDQKNQNN SPYHDACPATQQPSVIQAEKNKPQPLPTEGAE STHQAAHIQLSNPSSLSNIDFYAQVSDITPAGS VVLSPGQKNKAGMSQCDMHPEMVSLCQENF LMDNAYFCEADAKKCIPVAPHIKVESHIQPYS LNQEDIYITTESLTYTAAGSPGTGEHVPGSEM

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
458	1808	A	3663	sequence	462	nucleotide insertion PVPDYTSIHIVQSPQGLILNATALPLPDKEFLS SCGYVSTDQLNKIMP TRAPASGRSGAGLALSANAPDSGGHPGATEG
430	1000	^	3003	134	402	PAGSLAHASGSARGTWRVRGRGSHGWERTV GAGGCANPVPALHSCASAPRGTGRVSALGPK TGSSPLSSPKG
459	1809	A	3664	902	135	LGKYNTSMALFDFVLHNSTGEIRYITEDDVIQ SQNALGKYNTSMALFESNSFEKTILESPYYVD LNQTLFVQVSLHTSDPNLVVFLDTCRASPTSD FASPTYDLIKSGCSRDETCKVYPLFGHYGRF QFNAFKFLRSMSSVYLQCKVLICDSSDHQSRC \NQGCVSRSKRDISSYKWKTDSIIGPIRLKRDR SA\NGNSGFQHETHAEETPNQPFNSVHLFSFM VLALNVVTVATITVRHFVNQRADYQ\YQKLQ NY
460	1810	A	3670	850	557	LGILMSPQVEAGEI*ALLTPPPGCMQFSPLTL/P K*WVSPGLTP/PPPEVPSVFLVEPGLPHAGQA GLDLL\TSGDPPASTSQSARTTDVSHRAQPLAI .S
461	1811	A	3671	2472	2099	IGVLAFETGSCSVTRLYCIGIIMPHCSLDLAGS\ TSAFRIAGTTSVHHHPQLTFFFFWIETGSHCV VQTGL*LLALSNPPALASQIAGISGMSHRAWP GLVLYSLEFSLLCASQSLIMLFTCYNE
462	1812	Α	3672	394	110	VKPVNGESKRD*GADTQTCEGEADEQLQT\N CYYD/STKSFFYISCG*K\RKPTWAENRRLNA KMFGIPLHSNSDPWGYEEREVIGFHRSRVSRG HGS
463	1813	A	3673	348	1	QRNPFSAGHPQRPPTSGSQSELLAQPRLRPGR KSSFSRDQDVW*SQAVPKRQ*QRNPFSAGHP QRPPTSGSQSELLAQPRLRPGRKSSFSRDQDV WPGQKPRPSQQQHQMCASPTLGQRSPFALEP VPAYHGGRDPFASARPSPVGIPKPRAAPAGG GWRRIRPKSSTK
464	1814		3676	2253	320	PVIQRCSQPYGFSLLISFFLKCVSETSQQPPSR KVFQLLPSFPTLTRSKSHESQLGNRIDDVSSM RFDLSHGSPQMVRRDIGLSVTHRFSTKSWLS QVCHVCQKSMIFGVKCKHCRLKCHNKCTKE APACRISFLPLTRLRRTESVPSDINNPVDRAAE PHFGTLPKALTKKEHPPAMNHLDSSSNPSSTT FSTPSSPAPFPTSSNPSSATTPP\NPSP\GQR\DSR FNFPSC/AYFIHHR\Q\QFIFPDISAFAHAAPLPE AADGTRLDDQPKADVLEAHEAEAEEPEAGK SEAEDDEDEVDDLPSSRPWRGPISRKASQTS VYLQEWDIPFEQVELGEPIGQGRWGRVHRGR WHGEVAIRLLEMDGHNQDHLKLFKKEVMN YRQTRHENVVLFMGACMNPPHLAIITSFCKG RTLHSFVRDPKTSLDINKTRQIAQEIIKGMGY LHAKGIVHKDLKSRNVFYDNGKVVITDFGLF \(GISGVVP\EGRRENQLKLSHDWLCYLAPEIVR EMTPGKDEDQLPFSKAADVYAFGTVWYELQ ARDWPLKNQAAEASIWQIGSGEGMKRVLTS VSLGKEVSENLSACWAFDLQERPS\FSLLMD MLEKLPKLNRRLSHPGHF*KSADINSSKVVPR FERFGLGVLESSNPKM
465	1815	A	3679	8	803	IPSPAWWNSTWADTFSLLLALAVALYLGYY WACVLQTHRAFCASNTEDLETVVNHIKHRYP QAPLLAVGISFGGILVLNHILAQARQAAGLVA ALTLSACWDSFETTRSLETPLNSLLFNQPLTA GLCQLVERLSY/E*DLQARTIRQFDERYTSVA

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion FGYQDCVTYYKAASPRTKIDAIRIPVLYLSAA DDPFSTVCALPKQAAQHSPYVALLITARGGHI
466	1816	A _.	3684	3	307	GFLEGLLPWQHWYMSRLLHQYAKAIFQDPE GLPDLRALLPSEDRNS SSQYIVQSKTKIFL*AAREKQ/RHTCRRFSIRLS ANISSQTGEARGQWPSVFKVLKEKKLSTKKS FGQK*GR\RKTFPDKQK/LREFDTTRPTIQEML
467	1817	A	3687	2465	837	TGVLQG ELPTPLIAAHQLYNYVADHASSYHMKPLRMA RPGGPEHNEYALVSAWHSSGSYLDSEGLRHQ DDFDVSLLVCHCAAPFEEQGEAERHVLRLQF FVVLTSQRELFPRLTADMRFRKPPRLPPEPE APGSSAGSPGEASGLILAPGPAPLFPPLAAEVG MARARLAQLVRLAGGHCRRDTLWKRLFLLE PPGPDRLRLGGRLALAELEELLEAVHAKSIGD IDPQLDCFLSMTVSWYQSLIKVLLSRFPQSCR HFQSPDLGTQYLVVLNQKFTDCFVLVFLDSH LGKTSLTVVFREPFPVQPQDSESPPAQLVSTY HHLESVINTACFTLWTRLL*GSGLDH*MSLFL ESWAYQIACQRQD*PALLGPRASQTLSDTKG FVTMS*GSAAPAWQQEPPSPNTHSH*PIQDSR ESGQPRGPLGPFWGTPFGPPGRVSGVHTGWQ TPPRAPLPESCPLVPLTTVSHLCPLSLRVFTSHL DITAGHSHRDDTWVPIPALPLKHLRPPSSPFA LGPWVSHPLMRWVQKLSHLHSNPGTGFSMG GKQORN
468	1818	A	3691	960	499	QTCRKDKRAIYPHFQNE*MNEIKAI*SGTGGI QCFHSQNDSAFFFFLFLLETEFCSAA/TVQWH DFLSMQPPPPGFKQFTCLSLLSSWNYRR\PPPF PGNF*FLVKTGFPHVGQTGFELLTSSDLAPLA SQNGGITGMSPCAWPFFFFFFGLC
469	1819	A	3714	4747	495	MAYSWQTDPNPNESHEKQYEHQEFLFVNQP HSSSQVSLGFDQIVDEISGKIPHYESEIDENTFF VPTAPKWDSTGHSLNEAHQISLNEFTSKSREL SWHQVSKAPAIGFSPSVLPKPQNTNKECSWG SPIGKHHGADDSRFSILAPSFTSLDKINLEKEL ENENHNYHIGFESSIPPTNSSFSSDFMPKEENK RSGHVNIVEPSLMLLKGSLQPGMWESTWQK NIESIGCSIQLVEVPQSSNTSLASFCNKVKKIR ERYHAADVNFNSGKIWSTTTAFPYQLFSKTK FNIHIFIDNSTQPLHFMPCANYLVKDLIAEILH FCTNDQLLPKDHILSVWGSEEFLQNDHCLGS HKMFQKDKSVIQLHLQKSREAPGKLSRKHEE DHSQFYLNQLLEFMHIWKVSRQCLLTLIRKY DFHLKYLLKTQENVYNIEEVKKICSVLGCVE TKQITDAVNELSLILQRKGENFYQSSETSAKG LIEKVTTELSTSIYQLINVYCNSFYADFQPVNV PRCTSYLNPGLPSHLSFTVYAAHNIPETWVHR INFPLEIKSLPRESMLTVKLFGIACATNNANLL AWTCLPLFPKEKSILGSMLFSMTLQSEPPVEM ITPGVWDVSQPSPVTLQIDFPATGWEYMKPD SEENRSNLEEPLKECIKHIARLSQKQTPLLLSE EKKRYLWFYRFYCNNENCSLPLVLGSAPGW DERTVSEMHTILRRWTFSQPLEALGLLTSSFP DQEIRKVAVQQLDNLLNDELLEYLPQLVQAV KFEWNLESPLVQLLLHRSLQSIQVAHRLYWL LKNAENEAYFKSWYQKLLAALQFCAGKALN DEFSKEQKLIKILGDIGERVKSASDHQRQEVL KKEIGRLEEFFQDVNTCHLPLNPALCIKGIDH DACSYFTSNALPLKITFINANLMGKNISIIFKA

CEC ID	CEOID	Met	CEO	Dradiated	Dendiated and	Amino soid assumes (A-Alanias C-Custains
SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid	Predicted end nucleotide location corresponding to last amino acid residue of peptide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
				residue of peptide sequence	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						GDDLRQDMLVLQLIQVMDNIWLQEGLDMQ MIIYRCLSTGKDQRLVQMVPDAVTLAKIHRH SGLIGPLKENTIKKWFSQHNHLKADYEKALR NFFYSCAGWCVVTFILGVCDRHNDNIMLTKS
						GHMFHIDFGKFLGHAQTFGGIKRDRAPFIFTS EM'EYFITEGG'KNPQHFQDFV'ELCCRAYNIIR KHSQLLL'NLL'EMMLYAG'LPELSGI'QDLKY VYNNLRPQDTDLEATSHFTKKIKESLECFPVK
						LNNLIHTLAQMSAISPAKSTSQTFPQESCLLST TRSIERATILGFSKKSSNLYLIQVTHSNNETSL TEKSFEQFSKLHSQLQKQFASLTLPEFPHWW HLPFTNSDHRRFRDLNHYMEQILNVSHEVTN
						SDCVLSFFLSEAGQQTVEESSPVYLGEKFPDK KPKVQLVISYEDVKLTILVKHMKNIHLPDGSA PSAHVEFYLLPYPSEVRRRKTKSVPKCTDPTY NEIVVYDEVTELQGHVLMLIVKSKTVFVGAI NIRLCSVPLDKEKWYPLGNSII*PLLLFSSFGM
[i		[KSLEKDEFVGGMLLSNPIW
470	1820	A	3718	430	75	SHGSISILNLHQGCVFLPSLPAQGLRCYRCLA VLEGASCSVVSCPFLDGVCVSQKVSV/CWQ*/ CPWGARAEGRLSAVVDSQISCCKGDLCNAV
471	1821	A	3723	891	494	VLAAGSPWALCVQLLLSLGSVFLWALL LRQSL/NSVPQAGVQWRDSSLQAPPPRFTPLS CLSLPSSWDYRRLPPCLANFLYF**RRGFTML ARMVLIS*PRDPPASASQ\STEITGGSHRAQHP
			·			TDSRDHSERSVKKSHEVISELRMKVIKCKVAF
472	1822	A	3734	443	251	SKNPI GFIET*NFCVSKDTSKKLS/RLPTKWKNVFAN *ISDKGLVSRICQELLRHLDAEQVSSTAGLSL
473	1823	A	3746	.3	500	THASGGARSGAGWAGRGVRAGTEAGRGGIF LTLSILRTRDLPSGAMSEGVDLIDIYADEEFNQ DPEFNNTDQIDLYDDVLTATSQPSDDRSSSTE PPPPVRQEPSPKPNNKTPAILYTYSGLRNRRA AVYVGSFSWWTTDQQLIQVIRSIGVYDVGEV
474	1824	A	3753	2	5262	KFAENRAK RPLFAREGGIYAVLVCMQEYKTSVILVQQAG
						LAALKMLAVASSSEIPTFVTGRDSIHSLFDAQ MTREIFASIDSATRPGSESLLLTVPAAVILMLN TEGCSSAARNGLLLLNLLCNHHTLGDQIITQ ELRDTLFRHSGIAPRTEPMPTTRTILMMLLNR YSEPPGSP\ERAALETPIIQGQDGSPELLIRSLV
						GGPSAELLLDLERVLCREGSPGGAVRPLLKRL QQETQPFLLLLRTLDAPGPNKTLLLSVLRVIT RLLDFPEAMVLPWHEVLEPCLNCLSGPSSDSE IVQELTCFLHRLASMHKDYAVVLCCLGAKEI
				,		LSKVLDKHSAQLLLGCELRDLVTECEKYAQL YSNLTSSILAGCIQMVLGQIEDHRRTHQPINIP FFDVFLRHLCQGSSVEVKEDKCWEKVEVSSN PHRASKLTDHNPKTYWESNGSTGSHYITLHM
						HRGVLVRQLTLLVASEDSSYMPARVVVFGG DSTSCIGTELNTVNVMPSASRVILLENLNRFW PIIQIRIKRCQQGGIDTRVRGVEVLGPKPTFWP
						LFREQLCRRTCLFYTIRAQAWSRDIAEDHRRI. LQLCPRLNRVLRHEQNFADRFLPDDEAAQAL GKTCWEALVSPLVQNITSPDAEGVSALGWLL DQYLEQRETSRNPLSRAASFASRVRRLCHLL
						VHVEPPPGPSPEPSTRPFSKNSKGRDRSPAPSP VLPSSSLRNITQCWLSVVQEQVSRFLAAAWR APDFVPRYCKLYEHLQRAGSELFGPRAAFML

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						ALRSGFSGALLQQSFLTAAHMSEQFARYIDQ QIQGGLIGGAPGVEMLGQLQRHLEPIMVLSG LELATTFEHFYQHYMADRLLSFGSSWLEGAV LEQIGLCFPNRLPQLMLQSLSTSEELQRQFHLF QLQRLDKLFLEQEDEEEKRL*EEEEEEEEEA EKELFIEDPSPAISILVLSPRCWPVSPLCYLYHP RKCLPTEFCDALDRFSSFYSQSQNHPVLDMG PHRRLQWTWLGRAELQFGKQILHVSTVQMW LLLKFNQTEEVSVETLLKDSDLSPELLLQALV PLTSGNGPLTLHEGQDFPHGGVLRLHEPGPQ RSGEALWLIPPQAYLNVEKDEGRTLEQKRNL LSCLLVRILKAHGEKGLHIDQLVCLVLEAWQ KGPNPPGTLGHTVAGGVACTSTDVLSCILHLL GQGYVKRRDDRPQILMYAAPEPMGPCRGQA DVPFCGSQSETSKPSPEAVATLASLQLPAGRT MSPQEVEGLMKQTVRQVQETLNLEPDVAQH LLAHSHWGAEQLLQSYSEDPELLLAAGLCV HQAQAVPVRPDHCPVCVSPLGCDDDLPSLCC MHYCCKSCWNEYLTTRIEQNLVLNCTCPIAD CPAQPTGAFIRAIVSSPEVISKYEKALLRGYVE SCSNLTWCTNPQGCDRILCRQGLGCGTTCSK CGWASCFNCSFPEAHYPASCGHMSQWVDDG GYYDGMSVEAQSKHLAKLISKRCPSCQAPIE KNEGCLHMTCAKCNHGFCWRCLKSWKPNH KDYYNCSAMVSKAARQEKRFQDYNERCTFH HQAREFAVNLRNRVSAIHEVPPPRSFTFLNDA CQGLEQARKVLAYACVYSFYSQDAEYMDVV EQQTENLELHTNALQILLEETLLRCRDLASSL RLLRADCLSTGMELLRRIQERLLAILQHSAQD
ļ						FRVGLQSPSVEAWEAKGPNMPGSQPQASSGP EAEEEEEDDEDDVPEWQQDEFDEELDNDSFS
475	1825	A	3754	1093	96	YDESENLDQETFFFGDEEEDEDEAYD GTSRNQHSPKTHA*RSS/WPQPPPLFLPPLQPQ ATGRRRRTRTQQRTAALLTDGTTKTGAAW SRRPSLCWPSRTTGAPGAK*AVLVRSATPTTN PPNPQSPTGAAGKLRAPGNRAG/SEPSSQEPPP DGTR\RPASITGVAQSPATRATPSLPCLHVPAP SRGQTLGVRTTGRASRLTVDRSRLSWPGRSA RSGGGRWRPNAPRGRWPRAP*SWEPGSWTE PWRWPFPAAESPPHRCIYCTNHVSPAGPARPS HVYIIRATINSISHPLCRAQSSPWEAAGVWRR PAQPAPTSDVNINLLRKPRVKRHDLIYQFLGN TLWEEGRQRPPETLQPAR
476	1826	A	3758	901	521 ;	FFFGNGVSPCPQAGV*WHDLDSLQNLPPGFK RFSYLSLPSSW\DYRHVPPRQANFCIF/M*RRG FTMLARMVSIS*PRDLPALASQSAGITGVSHH APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI
477	1827	A	3761	843	575	GVISAHCNLRL/CHLPGSSNSPASASQVAGTIG ARTTPS*IFVFLVETGFHHVSQDGLDLL/NFVI RPRRPLKVLGLQACTRARLPSPLKEL
478	1828	A .	3763		1240	HLLSFHLWSASLDCLEQLSQERHVKGMLLGP PPVNESTKPSPSPWKLTPPMCSIPPVFPPKSGS PTTSWS/PSGHSKLEVERAQTGPFCLHIYCP*P GVTDNTTSLLHYIPFPRL\SGLVCFPAH*FPSY WTGHSFASQAWLRQVPEVSKHLQCPSAESLL TMEYHQPEDPAPGKAGTAEAVIPENHEVLAG PDEHPQDTDARDADGEAREREP/RRPSFAA*P VWGQP\ESPLPEASSAPPGPTLGTLPEVETIRA CSMPQELP*SPRTRQPEPDFYCVKWIPWKGE QTPHTQSTNGPLPSPCHHEHPLSSVEGEAPPA

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion EGSDHIG
479	1829	A	3766	2	2152	YSPIRLLEVCVPLPKIFIKRQAPLKVSLLQDLK DFFQKVSQVYVAIDERLASLKTDTFSKTREEK MEDIFAQKEMEEGEFKNWIEKMQARLMSSS VDTPQQLQSVFESLIAKKQSLCEVLQAWNNR LQDLFQQEKGRKRPSVPPSPGRLRQGEESKIS AMDASPRNISPGLQNGEKEDRFLTTLSSQSST SSTHLQLPTPPEVMSEQSVGGPPELDTASSSE DVFDGHLLGSTDSQVKEKSTMKAIFANLLPG NSYNPIPFPFDPDKHYLMYEHERVPIAVCEKE PSSIIAFALSCKEYRNALEELSKATQWNSAEE GLPTNSTSDSRPKSSSPIRLPEMSGGQTNRTTE TEPQPTKKASGMLSFFRGTAGKSPDLSSQKRE TLRGADSAYYQVGQTGKEGTENQGVEPQDE VDGGDTQKKQLINPHVELQFSDANAKFYCRL YYAGEFHKMREVILDSSEEDFIRSLSHSSPWQ ARGGKSGAAFYATEDDRFILKQMPRLEVQSF LDFAPHYFNYITNAVQQKRPTALAKILGVYRI GYKNSQNNTEKKLDLLVMENLFYGRKMAQ VFDLKGSLRNRNVKTDTGKESCDVVLLDENL LKMVRDNPLYIRSHSKAVLRTSIHSDSHFLSS HLIIDYSLLVGRDDTSNELVVGIIDYIRTFTWD KKLEMVVKSTGILGGQG*MPTVVSPELYRTR FCEAMDNYFLMVPDHCTGLGLNC
480	1830	A	3777	251	3	QGCGSAGTLIHY**ECKMVQLLWKTV*QFLI KLNI\KDPAITLDVYPNEVKNYVRTKTYTQMF I/ANFIMAKSWKQPTHPSVRT
481	1831	A _.	3779	333	3	EAAIRQPEPNILDVNQIFKDLAMIIHDQGDLID SIEANAESSEVLVERAPGQLQRPA\YYQKKSR KKMCLVVLVQTAIILICERIM*VVYTTKWSPPI VLPVSCFQGQKFN
482	1832	A	3780	2	371	TGGRQGKNDHTSITEKPSRDFNRHLITQNI*M PNQDMKSSSNSLIIRKVQIKPTILYHHIFTRKA KMKTTDKTKYR*GFKAITILIHCSQDCKLQ*S /L*ENHFMIFPKAEQHITYDTTIPFLR
483	1833	A	3787	43	448	LMKDLSPYVMETHYILNRLNER/RSMWRHIIG KLPNTKDQEKILKAIRGRREVIQGS/RQQYRR PAAFSAAEKARRLWCS/VFNIERRNL/CEYPTK LSFNIKGEMTFSDKTEFTTNRPSLKMLLKDRI QEEGKMF*KEKCFKRKE
484	1834	A	3798	1	727	FFFFETESRSVAQAGVQWCNLGSLQALPPGF\ SHSPASASRVAGTTGTRH*ARLIFYIFSRDGVS PC*PGWS*SPDLVIRPP\RLPKCWDYRREPPRP A*FFVFLVE\QGFTMLARMVSIS*PQ/CDLPAS VSQNAGITGVSHCAWPCLHFCFFGFFFEMESC SVAQAEVQWHDLRSLQAPPPGFTPFSCLSLPG SWDYRRPPPRPANF\CIFSRDGVSPC*PGWSRS PDLVIRPPRPPKVLGLQA
485	1835	A	3802	1	239	FFFFEMECLTVSQAGVQWYNLHSLQPLPPGF KQFSC\LSLPSSWD*RVPTSRPAKF/CVIF*DGV SHCQPGWSAVVQPPLH
486	1836	A	3811	378	98	RYD*SSQSENIP\QKEFLLKYP*CTATLGMRN MSIMKKKSIFSAEFYKVSLPSLLL\HLLAIEWG FHIEIQLTIHQHFLNYELESDFVHIVEYM
487	1837	A	3814	771	320	FDPDWTRAAGIRHEKKPKALAYRRENSPGDL PPPPLPPPEEEASWAL/GAEGSRQHVLPGAGA QWGEESGPGRAPGSPAGAPPR*RGLAP\NSRP SFLSRGQGTSTCSTAGSNSSRGSSSSRGSRGPG RSRSRSQSRSQSQRPGQKRREEPR

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
488	1838	A	3818		781	FRACLLELIPYAPTLSWTACPPAMAGPRGLLP LCLLAFCLAGFSFVRGQVLFKGCDVKTTFVT HVPCTSCAAIKKQTCPSGWLRELPDQITQDCR YEVQLGGSMVSMSGCRRKCRKQVVQKACCP GYWGSRCHECPGGAETPCNGHGTCLDGMDR NGTCVCQENFRGSACQECQDPNRFGPDCQSV CSCVHGVCNHGPRGDGSCLCFAGYTGPHCD QELPVWQELGFPQNNPRLRKAPNCKCLPG*H RNGLIATPNPCRP
489	1839	A	3822	934	669	FFFSEMESRSVTRLECSGAISAHLRLLGSSNSP ASAS*VAGTIGACHHAQLIFVFLVETGFHHVG
490	1840	A	3825	79	9748	QDGLDLL/NLMIHPPRPPKVLGFQA GCQSCWPAWPRLRRRGPASAGARLGRKAPW GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRETAMATLEKLMKAFESL KSFQQQQQQQQQQQQQQQQQQQQPPPP PPPPPPPQLPQPPPQAQPLLPQPQPPPPPPPPP GPAVAEEPLHRPKKELSATKKDRVNHCLTIC ENIVAQSVRNSPEFQKLLGIAMELFLLCSDDA ESDVRMVADECLNKVIKALMDSNLPRLQLEL YKEIKKNGAPRSLRAALWRFAELAHLVRPQK CRPYLVNLLPCLTRTSKRPEESVQETLAAAVP KIMASFGNFANDNEIKVLLKAFIANLKSSSPTI RRTAAGSAVSICQHSRRTQYFYSWLLNVLLG LLVPVEDEHSTLLILGVLLTLRYLVPLLQQQV KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEL TLHHTQHQDHNVVTGALELLQQLFRTPPPEL LQTLTAVGGIGQLTAAKEESGGRSRSGSIVELI AGGGSSCSPVLSRKQKGKVLLGEEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTPG SAGHDITTEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEEATGILPDEASEA FRNSSMALQQAHLLKNMSHCRQPSDSSVDKF VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLVHCVRLLSASFLLTGGKNVLVPDRDVRV SVKALALSCVGAAVALHPESFFSKLYKVPLD

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first	Predicted end nucleotide location corresponding to last amino acid residue	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
				amino acid residue of peptide sequence	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFVLKQFEYIEVGQFRESEAIIPNIFF FLVLLSYERYHSKQIIGIPKIIQLCDGIMASGR KAVTHAIPALQPIVHDLFVLRGTNKADAGKE LETQKEVVVSMLLRLIQYHQVLEMFILVLQQ CHKENEDKWKRLSRQIADIILPMLAKQQMHI DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILAILRVLISQSTED IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQIKNLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLIHIFKS GMFRRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITTHPALVLLWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRRGALILFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD
						PVPSLSPATTGALISHEKLLLQINPERELGSMS YKLGQVSIHSVWLGNSITPLREEEWDEEEEEE ADAPAPSSPTSPVNSRKHRAGVDIHSCSQFL LELYSRWILPSSSARRTPAILISEVVRSSLVVS DLFTERNQFELMYVTLTELRRVHPSEDEILAQ YLVPATCKAAAVLGMDKAVAEPVSRLLESTL RSSHLPSRVGALHGVLYVLECDLLDDTAKQL IPVISDYLLSNLKGIAHCVNIHSQQHVLVMCA TAFYLIENYPLDVGPEFSASIIQMCGVMLSGS EESTPSIIYHCALRGLERLLLSEQLSRLDAESL VKLSVDRVNVHSPHRAMAALGLMLTCMYT GKEKVSPGRTSDPNPAAPDSESVIVAMERVS VLFDRIRKGFPCEARVVARILPQFLDDFFPPQ DIMMKVIGEFLSNQQPYPQFMATVVYKVFQT LHSTGQSSMVRDWVMLSLSNFTQRAPVAMA TWSLSCFFVSASTSPWVAAILPHVISRMGKLE QVDVNLFCLVATDFYRHQIEEELDRRAFQSV LEVVAAPGSPYHRLLTCLRNVHKVTTC
491	1841	A	3826	469	302	SNPPASASRVAGITGVHQHAWLIFVFLVEMEF HHVGQAVLKLLISGDLPVSASQSA
492	1842	Ā	3836	392	88	VAPSPMIMPDLYFYRDPEEIEKEE*AAAEK\EE FQSEWTAVV/P/EFTATQSEVADWFKDMQVP SVPIQQFPTEDWST*PIMNDWSATSTAQTTE WVRITTEWP

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino said sequence (AmAlonine C. Curtaine
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutarnic Acid,
nucl-	peptide	1100	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine.
1		ł		1		
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
[[{	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			ł	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
ľ		ł	i	peptide	l '	/=possible nucleotide deletion, \=possible
		1	ŀ	sequence	1	nucleotide insertion
493	1843	A	3838	19	380	TPSDMNRAFETDTQSIGEKNRSPSEPDYFERK
175	1043	^	3636	19	360	
	1	1		[1	KFKRS*EKAHIRYKIDQPEDIPLK\EFLCKHSK
	1					CTATLSMRNMSLMKKKCSFSEEF\LAFFPSLL
		<u> </u>	ļ	i	İ	VCHLLAIKLGFYIEIHLTTFNNTF
494	1844	Α	3845	2	352	FFFLRRSL/DSVAQAEAQWL\ELGLLQAPPPGF
1	•	l	l	1	1	KPISLP\GLPSSWDYGRPPPCPANFCIF/M*RRG
		l	ļ			FTVLARMVLIS*PCDPPTLASQGTAITGMSYH
1		}	i	İ		
495	1045	-	2045	1001		ARPQDIDFLYAHQGRCWFRLL
493	1845	Α	3847	1774	40	DIFFRRAKEGMGQDEAQFSVEMPLTGKAYL
1			1			WADKYRPRKPRFFNRVHTGFEWNKYNQTHY
}	ļ.	į.	1	l	ł	DFDNPPPKIVQGYKFNIFYPDLIDKRSTPEYFL
	i					EACADNKDFAILRFHAGPPYEDIAFKIVNREW
					Ì	EYSHRHGFRCQFANGIFQLWFHFKRYRYRR*
			İ			RPWGTAGRCPRGHSKGASVKLVVTPGPLSGL
1		ł	l	}	1	
1			ļ.	1		QGRGFTSHLRPHLSFARPQFPPI*KGGHH*AC
	ł	ļ	i	ļ	}	HGELRRHWDRLA*GPDATEGALGASFEHEG
	1					GQQPPADLTVQADTLHRPSARLGGAHRACPK
	1	J			ł	RRPHRVLWRWARGAWAWRCQAREKQETQG
1						QPCHITGHPLGREAEPAAAGAAPALAHRPPF
						ARTGSTE\PGPCWRPIRHCRRDPLWTPTLC\RD
						WPPTHPVLAGGVHFPAAG/IGGCVEVPVSVN
1	1			İ		
	1	1				VMGTKSH*AVLPPPPSTGPGGQGLPEGWGLE
		Ì				KGEGLPPGIPPPGLLTGPW\SMRPVTPSFAHIR
1	ĺ	ĺ			ĺ	TVAPSHSPFSGQEGRGPHGCHSPGR\SGP\AGR
ļ		!				LVLQHPTGTSPTEAKRKVPPGPPEGHPTSPVT
i	1		1			SPRPPTAPPRHPASSGNSSVCFSKKTCRWEKK
	J	J]		j	SFVLMELAYWQDRMFF
496	1846	Α	3849	830	442	AKSPLPLG*IQWR/NLGSLKLRLPGFK*FTCLG
1 470	1040	Α	3047	630	442	
			[[LLSSWDYRSLPPRPVNFCILVELGFHHVDQAG
						LKLLTSSALPALASQSAEITGMSHRIWPLPLLR
						RPPVIRIRAPPORLPFNLITSLKALSPNMATF
497	1847	Α	3859	2	393	ALRKTRRDGIARTGAQPAASWKGTNNYPWR
		[LEMAGRPGSQEQSKDRGTGSLPPPSQRPLGPS
	1					PEGAGPSPPPPGIPRGGGSSSSEGP/PQLLFVPR
i		[1 '			RFPAPKKGLPSDTPHSKAPPTPHLILGGEDSQ
1		ł	ļ i			
100						VPIL
498	1848	Α	3860	253	634	KNASTVYSSQGDPKSFFFLLRWSLALVAQAG
					-	EQ*RDLSSLQPPPPGFK*FSCLSLPSSWD\YRCP
1			[,		LPCLANF*FLVETGFHHVGQADLKLLTSGDP
1		[1		PTSASESAGITGVSHRAWPRIHFLYWKTFFL
499	1849	A	3863	423	263	APSQISVAFLYAA/DKLFEKEI*KKIPFIIAS/DKI
[i 5005		200	=
]						KIGINLTKEVKYLYTENYITLMKEIK/DTDKW
}		1		ļ		KDILY*WIGKINI*KMSTPPKAIYRFNAIPTKIP
1 .						MTFFTEIEKSIIKFIWNIIKKPPNTQSNIEQKE*S
				•		FCSILLWVFGGFLWFHMNFMIDFSISVKNVIGI
				l		LVGIALNL
500	1850	A	3865	2	15246	LPRGCLWCLQRSPTPARPQPSRPARSPLPLFP
				-		DLRPWASDLDIMGDAEGEDEVQFLRTDDEV
1				ļ	į	
						VLQCSATVLKEQLKLCLAAEGFGNRLCFLEP
						TSNAQNVPPDLAICCFVLEQSLSVRALQEML
] [ANTVEAGVESSQGGGHRTLLYGHAILLRHAH
1 1			1	ŀ	}	SRMYLSCLTTSRSMTDKLAFDVGLQEDATGE
				ļ	1	ACWWTMHPASKQRSEGEKVRVGDDIILVSVS
, ,				ļ		SERYLHLSTASGELQVDASFMQTLWNMNPIC
] [, ,		ļ	ļ		SRCEEGFVTGGHVLRLFHGHMDECLTISPADS
j l			1		1	
} [ĺ		[1		DDQRRLVYYEGGAVCTHARSLWRLEPLRIS
			ŀ	ŀ		WSGSHLRWGQPLRVRHVTTGQYLALTEDQG
			ł	ļ		LVVVDASKAHTKATSFCFRISKEKLDVAPKR
			1	ì		DVPCMCDDPHAYCEST OPTIONAL COLUMNS A
] 1	J	1	Į		ŀ	DVEGMGPPEIKYGESLCFVQHVASGLWLTYA

SEQ ID NO: of nucl- eotide scq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cystelne, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion APDPKALRLGVLKKKAMLHOEGHMDDALSL
						TRCQQEESQAARMIHSTNGLYNQFIKSLDSFS GKPRGSGPPAGTALPIEGVILSLQDLIIYFEPPS
]			EDLQHEEKQSKLRSLRNRQSLFQEEGMLSMV LNCIDRLNVYTTAAHFAEFAGEEAAESWKEI
						VNLLYELLASLIRGNRSNCALFSTNLDWLVS KLDRLEASSGILEVLYCVLIESPEVLNIIQENHI
						KSIISLLDKHGRNHKVLDVLCSLCVCNGVAV RSNQDLITENLLPGRELLLQTNLINYVTSIRPN
						IFVGRAEGTTQYSKWYFEVMVDEVTPFLTAQ ATHLRVGWALTEGYTPYPGAGEGWGGNGV
			!			GDDLYSYGFDGLHLWTGHVARPVTSPGQHL LAPEDVISCCLDLSVPSISFRINGCPVQGVFESF
						NLDGLFFPVVSFSAGVKVRFLLGGRHGEFKF LPPPGYAPCHEAVLPRERLHLEPIKEYRREGP
			1			RGPHLVGPSRCLSHTDFVPCPVDTVQIVLPPH LERIREKLAENIHELWALTRIEQGWTYGPVRD
						DNKRLHPCLVDFHSLPEPERNYNLQMSGETL KTLLALGCHVGMADEKAEDNLKKTKLPKTY
						MMSNGYKPAPLDLSHVRLTPAQTTLVDRLAE
			ļ			NGHNVWARDRVGQGWSYSAVQDIPARRNPR LVPYRLLDEATKRSNRDSLCQAVRTLLGYGY
						NIEPPDQEPSQVENQSRCDRVRIFRAEKSYTV QSGRWYFEFEAVTTGEMRVGWARPELRPDV
						ELGADELAYVFNGHRGQRWHLGSEPFGRPW QPGDVVGCMIDLTENTIIFTLNGEVLMSDSGS
			•			ETAFREIEIGDGFLPVCSLGPGQVGHLNLGQD VSSLRFFAICGLQEGFEPFAINMQRPVTTWFS
						KGLPQFEPVPLEHPHYEVSRVDGTVDTPPCLR LTHRTWGSQNSLVEMLFLRLSLPVQFHQHFR
						CTAGATPLAPPGLQPPAEDEARAAEPDPDYE NLRRSAGGWSEAENGKEGTAKEGAPGGTPQ
						AGGEAQPARAENEKDATTEKNKKRGFLFKA KKVAMMTQPPATPTLPRLPHDVVPADNRDD
		ļ				PEILNTTTYYYSVRVFAGQEPSCVWAGWVT PDYHQHDMSFDLSKVRVVTVTMGDEQGNV
	:					HSSLKCSNCYMVWGGDFVSPGQQGRISHTDL VIGCLVDLATGLMTFTANGKESNTFFQVEPN
						TKLFPAVFVLPTHQNVIQFELGKQKNIMPLSA AMFQSERKNPAPQCPPRLEMQMLMPVSWSR
						MPNHFLQVETRRAGERLGWAVQCQEPLTMM ALHIPEENRCMDILELSERLDLQRFHSHTLRL
						YRAVCALGNNRVAHALCSHVDQAQLLHALE DAHLPGPLRAGYYDLLISIHLESACRSRRSML
						SEYIVPLTPETRAITLFPPGRSTENGHPRHGLP GVGVTTSLRPPHHFSPPCFVAALPAAGAAEAP
						ARLSPAIPLEALRDKALRMLGEAVRDGGQHA RDPVGASVEFQFVPVLKLVSTLLVMGIFGDE
						DVKQILKMIEPEVFTEEEEEEDEEEGEEDEE EKEEDEETAQEKEDEEKEEEAAEGEKEEG
		ļ				LEEGLLQMKLPESVKLQMCHLLEYFCDQELQ HRVESLAAFAERYVDKLQANQRSRYGLLIKA
						FSMTAAETARRTREFRSPPQEQINMLLQFKDG TDEEDCPLPEEIRQDLLDFHQDLLAHCGIQLD
		ł	ŀ			GEEEEPEEETTLGSRLMSLLEKVRLVKKKEEK PEEERSAEESKPRSLQELVSHMVVRWAQEDF
						VQSPELVRAMFSLLHRQYDGLGELLRALPRA YTISPSSVEDTMSLLECLGQIRSLLIVQMGPQE
						ENLMIQSIGNIMNNKVFYQHPNLMRALGMHE TVMEVMVNVLGGGESKEIRFPKMVTSCCRFL

	SEQ ID NO: of nucl- eotide seq-	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496	Predicted beginning nucleotide location correspondi	Predicted end nucleotide location corresponding to last amino	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
	uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		· ·			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
					residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		İ		i	peptide		/=possible nucleotide deletion, \=possible
					sequence		nucleotide insertion
							CYFCRISRQNQRSMFDHLSYLLENSGIGLGM QGSTPLDVAAASVIDNNELALALQEQDLEKV
							VSYLAGCGLQSCPMLVAKGYPDIGWKPCGG
				ļ			ERYLDFLRFAVFVNGESVEENANVVVRLLIR
							KPECFGPALRGEGGSGLLAAIEEAIRISEDPAR
							DGPGIRRDRRREHFGEEPPEENRVHLGHAIMS
							FYAALIDLLGRCAPEMHLIQAGKGEALRIRAI
						•	LRSLVPLEDLVGIISLPLQIPTLGKDGALVQPK
					·		MSASFVPDHKASMVLFLDRVYGIENQDFLLH VLDVGFLPDMRAAASLDTATFSTTEMALAV
						!	NRYLCLAVLPLITKCAPLFAGTEHRAIMVDS
						•	MLHTVYRLSRGRSLTKAQRDVIEDCLMSLCR
							YIRPSMLQHLLRRLVFDVPILNEFAKMPLKLL
ı							TNHYERCWKYYCLPTGWANFGVTSEEELHL
					i		TRKLFWGIFDSLAHKKYDPELYRMAMPCLC
							AIAGALPPDYVDASYSSKAEKKATVDAEGNF DPRPVETLNVIIPEKLDSFINKFAEYTHEKWAF
							DKIQNNWSYGENIDEELKTHPMLRPYKTFSE
							KDKEIYRWPIKESLKAMIAWEWTIEKAREGE
							EEKTEKKKTAKISQSAQTYDPREGYNPQPPDL
i							SAVTLSRELQAMAEQLAENYHNTWGRKKKQ
						:	ELEAKGGGTHPLLVPYDTLTAKEKARDREKA
ļ							QELLKFLQMNGYAVTRGLKDMELDSSSIEKR FAFGFLQQLLRWMDISQEFIAHLEAVVSSGRV
							EKSPHEQEIKFFAKILLPLINQYFTNHCLYFLS
				ĺ			TPAKVLGSGGHASNKEKEMITSLFCKLAALV
							RHRVSLFGTDAPAVVNCLHILARSLDARTVM
						:	KSGPEIVKAGLRSFFESASEDIEKMVENLRLG
							KVSQARTQVKGVGQNLTYTTVALLPVLTTLF QHIAQHQFGDDVILDDVQVSCYRTLCSIYSLG
ı		·					TKNTYVEKLRPALGECLARLAAAMPVAFLE
							PQLNEYNACSVYTTKSPRERAILGLPNSVEEM
		·					CPDIPVLERLMADIGGLAESGARYTEMPHVIE
		·					ITLPMLCSYLPRWWERGPEAPPSALPAGAPPP
-							CTAVTSDHLNSLLGNILRIIVNNLGIDEASWM KRLAVFAQPIVSRARPELLQSHFIPTIGRLRKR
1							AGKVVSEEEQLALEAKAEAQEGELLVRDEFS
							VLCRDLYALYPLLIRYVDNNRAQWLTEPNPS
							AEELFRMVGEIFIYWSKSHNFKREEQNFVVQ
						ن	NEINNMSFLTADNKSKMAKAGDIQSGGSDQE
		. ј				ļ	RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN MCAPTDQDLITLAKTRYALKDTDEEVREFLH
1							NNLHLQGKVEGSPSLRWQMALYRGVPGREE
	-			ļ			DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS
			1		i		KKAVWHKLLSKQRRRAVVACFRMTPLYNLP
						.	THRACNMFLESYKAAWILTEDHSFEDRMIDD
	.		ĺ	[18	LSKAGEQEEEEEEVEEKKPDPLHQLVLHFSRT ALTEKSKLDEDYLYMAYADIMAKSCHLEEG
				į		1	GENGEAEEEVEVSFEEKQMEKQRLLYQQARL
-				İ	- 1	1	HTRGAAEMVLQMISACKGETGAMVSSTLKL
			l	İ		~	GISILNGGNAEVQQKMLDYLKDKKEVGFFQS
1	į		İ	1		Ì	IQALMQTCSVLDLNAFERQNKAEGLGMVNE
	į	1			į		DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTTINIICTVDYLL
	ŀ	j	1		į		RLQESISDFYWYYSGKDVIEEQGKRNFSKAM
-	ļ	ļ			İ		SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW
-		}	}	}		ł	DAVVGFLHVFAHMMMKLAQDSSQIELLKEL
1							LDLQKDMVVMLLSLLEGNVVNGMIARQMV
		ļ					DMLVESSSNVEMILKFFDMFLKLKDIVGSEAF
L							QDYVTDPRGLISKKDFQKAMDSQKQFSGPEI

SEQ ID NO: of nucl- eotide seq-	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496	Predicted beginning nucleotide location correspondi	Predicted end nucleotide location corresponding to last amino	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline.
uence			914	ng to first amino acid residue of peptide sequence	acid residue of peptide sequence	Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				Sequence		QFLLSCSEADENEMINCEEFANRFQEPARDIG FNVAVLLTNLSEHVPHDPRLHNFLELAESILE YFRPYLGRIEIMGASRRIERIYFEISETNRAQW EMPQVKESKRQFIFDVVNEGGEAEKMELFVS
						FCEDTIFEMQIAAQISEPEGEPETDEDEGAGA AEAGAEGAEEGAAGLEGTAATAAAGATARV VAAAGRALRGLSYRSLRRRVRRLRRLTAREA ATAVAALLWAAVTRAGAAGAGAAAGALGL LWGSLFGGGLVEGAKKVTVTELLAGMPDPT
						SDEVHGEQPAGPGGDADGEGASEGAGDAAE GAGDEEEAVHEAGPGGADGAVAVTDGGPFR PEGAGGLGDMGDTTPAEPPTPFGSPILKRKLG VDGVEEELPPEPEPEPEPELEPEKADAENGEK
		:		·		EEVPEPTPEPPKKQAPPSPPPKKEEAGGEFWG ELEVQRVKFLNYLSRNFYTLRFLALFLAFAIN FILLFYKVSDSPPGEDDMEGSAAGDVSGAGS GGSSGWGLGAGEEAEGDEDENMVYYFLEES TGYMEPALRCLSLLHTLVAFLCIIGYNCLKVP
						LVIFKREKELARKLEFDGLYTTEQPEDDDVKG QWDRLVLNTPSFPSNYWDKFVKRKVLDKHG DIYGRERIAELLGMDLATLEITAHNERKPNPP PGLLTWLMSIDVKYQIWKFGVIFTDNSFLYLG
	·			•		WYMVMSLLGHYNNFFFAAHLLDIAMGVKTL RTILSSVTHNGKQLVMTVGLLAVVVYLYTVV AFNFFRKFYNKSEDEDEPDMKCDDMMTCYL FHMYVGVRAGGGIGDEIEDPAGDEYELYRVV FDITFFFFVIVILLAIIQGLIIDAFGELRDQQEQV
501	1851	A	3869	467	665	KEDMETKCFICGIGSDYFDTTPHGFETHTLEE HNLANYMFFLMYLINKDETEHTGQESYVWK MYQERCWDFFPAGDCFRKQYEDQLS VIVAIYCQLIFDKGAKTIQ*PFQQIAL/CKRMK
						LGPCFTPCGKINSEWIRELSVRVKTIKHLEIGV N
502	1852	A	3888	1042	724	SGMQWRDLTPLQPLPPRFKQFSCLSLPGSWD YRHAP\PLLTNF*FLVEMGFCYVGQAGRKLL ASSDQSALASQSAGITGISTAPGPPFFFLNFEA GSCSVAQAGVQ
503	1853	A	3891	1773	1193	EVDSQSGVQ*QAPGSLQLQTPGLK/VSCLLSR QDYRSSLPHLASCCYYYYYY/VFL*RRGLTTL VQGGLKLLPSSNPFASAP*TAGITGMSHCAGP HFNF*MFRKISCIRE*F*HTRIYDIPFLILFFKET WVLLCYPGWPQIPGLKPSSCLRLLSSWDHRC APPCPASFFIFHVDRVSPPCPGLVSITFKMLLL L
504	1854	В	3896	279	70	MVSKSKSILMSYNHVELTFSDMKKMPEAFRR TQKHTIYLIPYQVIFWSTGKDAMRSFMMPFY QKEYYENQ*
505	1855	A	3899	2	1396	EPGVPTKKTWFDKPDFNRTNSPGFQKKVQFG NENTKLELRKVPPELNNISKLNEHFSRFGTLV NLQVAYNGDPEGALIQFATYEEAKKAISSTEA VLNNRFIKVYWHREGSTQQLQTTSPKVMQPL VQQPILPVVKQSVKERLGPVPSSTIEPAEAQS ASSDLPQVLST\LLA*QKQCIIQLL/WKAAQKT LLVSTSAVDNNEAQKKKQEALKLQQDVRKR KQEILEKHIETQKMLISKLEKNKTMKSEDKAE IMKTLEVLTKNITKLKDEVKAASPGRCLPKSI KTKTQMQKELLDTELDLYKKMQAGEEVTEL RRKYTELQLEAAKRGILSSGRGRGIHSRGRGA VHGRGRGRGRGRGRGVPGHAVVDHRPRALEIS

CEO ID	CEO ID	Mat	Tero	l Desires		
SEQ ID NO: of	SEQ ID NO: of	Met	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	1100	in NO.	nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1			Ì	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	Soquence	/=possible nucleotide deletion, \=possible
	İ	1		sequence		nucleotide insertion
		1	 	1 1		AFTESDREDLLPHFAQYGEIEDCQIDDSSLHA
		1		ļ	}	VITFKTRAEAEAAAVHGARFKGQDLKLAWN
	j			Ì		KPVTNISAVETEEVEPDEEEQREIIIA
506	1856	A	3911	1952	919	DAELSGTLSLVLTQCCKRIKDTVQKLASDHK
			1	ļ		DIHSSVSRVGKAIDKNFDSDISSVGIDGCWQA
ļ		}				DSQRLLNEVMVEHFFRQGMLDVAEELCQES
1	1	1		!		GLSVDPSQKEPFVELNRILEALKVRVLRPALE
1		j		}		WAVSNREMLIAQNSSLEFKLHRLYFISLLMG
			1			GTTNQREALQYAKNFQPFALNHQKDIOVLM
		1	1	İ		GSLVYLRQGIENSPYVHLLDANQWADICDIFT
				!		RDACALLGLSVESPLSVSFSAGCVALPALINIK
	j				•	AVIEQRQCTGVWNQKDELPIEV\DLG*KSAGY
1	İ	1	1	ł		HSIFACPILRQQTTDNNPPMKLVCGHIISRDAL
L	ļ	1				NKMFNGSKLKCPYCPMEQSPGDAKQIFF
507	1857	A	3936	439	18	SHPFSPAPGICPDAPPPLPRPSKGLGHPGTAGA
1		l		i i		PGSGARCHPPSTCSPSWASPG*GAKASPALPR
						SHGVTLLCKAQAHLCRGEDSKDASGSTSOA
1		İ	ł			WEPG*GAWGMPRCQGPALGSCFCPPGTTVQ
						RPAKORDKRNRHLGR
508	1858	Α	3944	120	412	WCPAGTLDFPGPQEMVLLEIEVMNQLNHRNL
Í	ĺ	ĺ		1		IQLYAAIETPHEIVLFME\YECPK*W*GLGGGT
	1	1				TRHGASRGGVCAHSIEGGELFERIVDEDYHLT
						EV
509	1859	A	3949	31	392	LTKTPSPREKGRGVLSVLLMMI*KCRVIFVKIP
i		ļ			ļ·	MVFFLQNFC/RIILNVA\WTGD*PNTL*KEQRG
						ITFSDSKS*YKATKIKTMWYCHKNRYID/ERN
		L				RIEIPEINPCICDKIIFRKLSMTTQ
510	1860	Α	3954	1013	885	FSETRACCPRLEHSGRIEAHCSLNIPGSSDPPT
						SASSVAATTG
311	1861	- A	3956	1	1054	PPAWAPRSPLIWAPTSGRHPCRAALPWSTSSV
1	İ					RWQPSEKQPPPPAHRGPADSLSTAAGAAELS
1		i				AEGAGKSRGSGEQDWVNRPKTVRDTLLALH
1	Ì	ľ				QHGHSGPFESKFKKEPALTAVARTARKRKPS
i						PEPEGEVGPPK\TTERPSRGCPHPQRGSRSP*L
						LHPLLCLRHHPLPHLIPTGPHRLKRPRM\P\SP
1	ļ]				MAALILVADNAGGSHASKDANQVHSTTRRN
1]	j				SNSPPSPSSMNQRRLGPREVGGQGAGNTGGL
		l				EPVHPASLPDSSLATSAPLCCTLCHERLEDTH
		1				FVQCPSVPSHKFCFPCSRQSIKQQGASGEVYC
						PSGEKCPLVGSNVPWAFMQGEIATILAGDVK
512	1062		2067	1000		VKKERDS
512	1862	Α	3957	1086	3	QDRARLDCSSATSAHCNLRLPGS*DSPASASR
						VAGTTDTHHHTWLILGSSVQTGFDHVGQAG
						LELLTSGDPPISASESAGIMGMSHCVWP*SWG
					J	LSHHMAPPQGDGGRARGTPGPEQSFWNLSC
						H*PRCQVPS*LMTQL/FWGRHQYNPTMKRGK
						LRHREACSLPLPGEGEPGLQPSS*SQNPCSSPL
						FHHGL*AWLWCPELLLQGQARRH*RSPPS/FK
						CPATLSLTAWSQTKRLRSQFLLLPWL*RAL*H
				1		PP\CHWPSRRSLGDPLLPRSQG*RDGT*ASTFC
						SYF*DTESHLVAQAGVQWRDLGSLQPPCPRL
		İ		ļ		K\RFSRLSPPSSYTHRYVPSHLAESCISSRDRIP
512	1060	L	206	2000		PSRPDRSRNSNSLSR
513	1863	Α	3961	3038	476	VALTTSMCCNKQVIVIDKIKSASIADRCGALH
				ĺ	Ĭ	VGDHILSIDGTSMEYCTLAEATQFLANTTDQ
					Į.	VKLEILPHHQTRLALKGPDHVKIQRSDRQLT
	İ		i	ı	ı	TREELES THIS TREAL ROLD IN A RIGHT PROPERTY OF THE PROPERTY OF
						WDSWASNHSSLHTNHHYNTYHPDHCRVPAL
						WDSWASNHSSLHTNHHYNTYHPDHCRVPAL TFPKAPPPNSPPALVSSSFSPTSMSAYSLSSLN MGTLPRSLYSTSPRGTMMRRRLKKKDFKSSL

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						SLASSTVGLAGQVVHTETTEVVLTADPVTGF GIQLQGSVFATETLSSPPLISYIEADSPAERCG VLQIGDRVMAINGIPTEDSTFEEASQLLRDSSI TSKVTLEIEFDVAESVIPSSGTFHVKLPKKHN VELGITISSPSSRKPGDPLVISDIKKGSVAHRT GTLELGDKLLAIDNIRLDNCSMEDAVQILQQC EDLVKLKIRKDEDNSDEQESSGAIIYTVELKR YGGPLGVITISGTEEPYDL*IISSLTKGGLAERT GAIHIGDRIL\AINSSSLKGKPLSEAIHLLQMAG ETVTLKIKKQTDAQSASSPKKFPISSHLSDLGD VEEDSSPAQKPGKLSDMYPSHGCPSVDSAVD SWDGSA\IDTS\YGTEGT\SFQASGY\NFNTYD WRSPKQRGS\LSPVT\KPRSQTYPDVGLSYED WDRSTASGFAGAA\IDSAETEQEENFWSQALE DLETCGQSGILRELEATIMSGSTMSLNHEAPT PRSPAGSDRPSFQERSSSRPHYSQTTRSNTLPS DVGRKSVTLRKMKQEIKEIMSPTPVELHKVT LYKDSDMEDFGFSVADGLLEKGVYVKNIRPA GPGDLGGLKPYDRLLQVNHVRTRDFDCCLV VPLIAESGNKLDLVISRNPLASQKSIDQQSLPG D*SEQNSAFFQQPSHGGNLETREPTNTL
514	1864	A	3967	833	800	LEKQĞVSGMATKRLARQLGLIRRKSIAPANG NLGRSKSKQLFDYLIVIDFESTCWNDGKHHH SQEIIEFPAVLLNTSTGQIDSEFQAYVQPQEHPI LSEFCMELTGIKQAQVDEGVPLKICLSQFCK WIHKIQQQKNIIFATGISEPS/DF*SKIMCICYL VR*RISYTY*SKHKSKGC
515	1865	A	3969	492	182	CRFWGISTHCDTCDPLSPQTTEG**EGDLWSL DLLGPEFLARKPLFKTKTYQSTF*SISKNE/FTC PNFIIEEGTDLIF*QVKHNPCHRLTPEEGTVQL NRADS
516	1866	A	3977	2	1357	KMLC/QKESNYIRLKRAKMDKSMFVKIKTLGI GAFGEVCLARKVDTKALYATKTLRKKDVLL RNQVAHVKAERDILAEADNEWVVRLYYSFQ DKDNLYFVMDYIPGGDMMSLLIRMGIFPESL ARFYIAELTCAVESVHKMGFIHRDIKPDNILID RDGHIKLTDFGLCTGFRWTHDSKYYQSGDHP RQDSMDFSNEWGDPSSCRCGDRLKPLERRAA RQHQRCLAHSLVGTPNYIAPEVLLRTGYTQL CDWWSVGVILFEMLVGQPPFLAQTPLETQM KVINWQTSLHIPPQAKLSPEASDLIIKLCRGPE DRLGKNGADEIKAHPIF*NQFDFSQ*PEDSRS AFKQFP*NHTTPTDTSNFDPJVDPDKLWSDDN EEENVNDTLNGWYKNGKHPEHAFYEFTFRRF FDDNGYPYNYPKPIEYEYINSQGSEQQSDEDD QNTGSEIKNRDLVYV
517	1867	A	3980	1358	1022	FFFKKFTQSLGFLLFSFSFLFSCFFFFHFVLFCY VFLDRVPLCHPGWSAVVQSQVT/VNLPPSWD *RCRPPH/LANLCNFCRD\SFTTLPRLVLNTWA QAIFQPQPPKVLGLQV
518	1868	A	3986	974	666	SPEMESHPITQAGVQWHHLSSLQPLPPGFK*F SCFSLPE*LGYRHVPPCLANSVFSVEMG\FLH VGQAGLELLTSGDLPALASQSAGITG\SHRAR PENGFENIF
519	1869	A	3994	751	126	NQGLRHVGLCRTCLVNQMFASSILGKSHHHS LISINQGHNALWKAAG\PLPLKAGYC\QSFSPC DSLKYG\SWDEKDLTVPQRDTHKRSVLRWIS QRGK\LAVEMEEGHCLL\LPLGTECLGIK\PIV HLFSSEMGE\NRPMVG\ARHVYSNAALLSFTP

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
	 	 		sequence		nucleotide insertion LRCLGGEKHKSGLHARPVIVPSLELHYDMDSI
520	1870	A	3999	882	698	AHV\FADLLLIITLPSYYIPFC QSFRLSLLSSWDYRHM*PRLANF*T\FFCRDR/
521	1871	A	4011	1346	1178	SLALLPRLVSNSWPQAILPPRPPKVLGLQT FFF*ETVSCSAS*AGVRSHDNSSLOPPSPG\SSN
522	1872	A	4015	2	377	PPTSASHVAGATGTHHHAWLLSV QGIALLTRMGESVKHVTGGYKLRTRPLEFAA
322	1072	A	4013		377	IGDYLDTFALKLGTIDRIAQRIIKEEIEYLVELR EYGPVYSTWSALEGELAEPLEGVSACIGNCST AL*ELTDDMTEDFLFVLREYILYSDSMK
523	1873	A	4018	341	19	ERVIHNQIQQAQRSPHIFNARRSS/PRPNIVELP KVKEVCKTSKS/GQVIYKGVSIRLRANFLAEP L*NRREWDEAIKVLKEKQ\FLSKMVYPANLSF GNEGDITSFPAK
524	1874	Α ·	4020	1067	743	FFLRWSL/DSVAQAGVKWCNLGSLQAPPPGF TPFSCLSLPSSWDYRHPPPRLAN*LTNFLCF** RQGFTVLARMVLIS*PHDLPASASQSAGITGL SHCSWPTSSILS
525	1875	Α	4021	781	351	QFRVIFFFLRRSHSVAQAGMQWHDHSLLQPL PPRLKQ/F/SHLSPPSIWDYRRVPPCLVNFSIFF VETGSCQPCLQLLGSSNPPASASQSAGIAGISH QGQPE*SFDIRFACVIAALRETFQCLCSASRVN NKIINRPTHPVESSF
526	1876	A	4024	80	341	TPSSTSRGTEEQQSSKMAWQRREEKEHLNVR RSSAEDGWKADKP/VDG*TPGEDHLPTPSPFQ LHIHSSESQLHHSVKSPPSLSFRLM
527	1877	Α	4026	593	230	DFYLYPERKKRGQMMTAVSLTTRPQESVAFE DVAVYFTTKEWAIMG\PAERALYRDVMLEN YGGCGPL*CHPTSKPALVFS\LEQGKESCFSPA TGSSLSRNDWRAGWIGYLELRRYTYLS
528	1878	Α	4028	1160	242	GTSELLCIQRWNWGPAFPPRPGLALAPTLQLL VEMGSAKSVPVTPARPPPHNKHLARVADPRS PSAGILRTPIQVESSPQPGLPAGEQLEGLKHAQ DSDPRSPTLGIARTPMKTSSGDPPSPLVKQLSE VFETEDSKSNLPPEPVLPPEAPLSSELDLPLGT QLS VEEQMPPWNQTEFPSKQVFSKEEARQPT ETPVASQSSDKPSRDPETPRSS\GSMRNRWKP\ NSSKVL\GKSPLHPSCQDDNSPGTLTLRQGKA AFKPLSENVSELK\EGA\ILGTGR\LLKTEGRA WEQGQD\HDKENQHFPLVES
529	1879	Α	4039	2	366	KDMVLIMEMQSMITMKCPQYL*E*RKIPDITK CW*GCGSTGILIFC/WS*PL*KTI*QPR*FKQI*T ILTIIYSIM*EHTFHNAGV*LSDIYPRFMKGYV HTEICT*MFIAVLFVVVKTWKQF
530	1880	Α	4057	358	3	LLEVNGNTIVTVFTKAQNKKNKGSRSILFKQL RKYGSRINLLKSKHDKNICTENYKT*MKEIEA /DTDKWKDILCSWIRRIHMKDILCSWIGRTHV VKISILPKVNYRFYLISIKIIMAI
531	1881	A	4061	50	278	TQGTEEIYKISSCEWVQASFSTPLITLHDFKIY HKATVIKMVWYWHRQ*KFSKN/RIESSEIEPH IYDQFIFDKGEKIIQEKGNSFFNN/MCWKNWIF T*KR
532	1882	A	4069	19	368	NDLLENFKFWE*FKE*LENINGTVTEKETGGV YKELSSPK YSGTROF YGQTISNFPGKIISMVY KLFQNTE/TEGRHPISL YEFRITLITIPNKDNIYL QIWMPVSLMNIVTLKCPT
533	1883	Α	4076	1	355	PIRKFTKVAG*KSNTPK*LAFLHINNEQFENKI/ ITNI/PFIIASKRIKYSGISLTKEMKDLYTETLLR KIKEDTNKWKDI/SCFWVGR/LNIVKMPK/VIC

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion IFNAIPIKMPMMCMAKIEKNSS
534	1884	A	4088	3	1931	IIIDSSTRRMESERSPLYRQLIDLGYLSSSHWNC GAPGQDTKAQSMLVEQSEKLRHLSTFSHQVL QTRLVDAAKALNLVHCHCLDIFINQAFDMQR DLQITPKRLEYTRKKENELYESLMNIANRKQE EMKDMIVETLNTMKEELLDDATNMEFKDVI VPENGEPVGTREIKCCIRQIQELIISRLNQAVA NKLISSVDYLRESFVGTLERCLQSLEKSQDVS VHITSNYLKQILNAAYHVEVTFHSGSSVTRM LWEQIKQIIQRITWVSPPAITLEWKRKVAQEAI ESLSASKLAKSICSQFRTRLNSSHEAFAASLRQ LEAGHSGRLEKTEDLWLRVRKDHAPRLARLS LESRSLQDVLLHRKPKLGQELGRGQYGVVYL CDNWGGHFPCALKSVVPPDEKHWNDLALEF HYMRSLPKHERLVDLHGSVIDYNYGGGSSIA VLLIMERLHRDLYTGLKAGLTLETRLQIALDV VEGIRFLHSQGLVHRDIKLKNVLLDKQNRAKI TDLGFCKPEAMMSGSIVGTPIHMAPELFTGK YDNSVDVYAFGILFWYICSGSVKLPEAFERCA SKDHLWNNVRRGARPERLPVFDEECWQLME ACWDGDPLKRPLLGIVQPMLQGIMNRLCKS\ NSEQPNRGLDDST
535	1885	Ā	4090	2	417	ALMPHEANYEEIFLKTDKDMDGFESGLEVRE IFLKTR/GLPSTLLAHIWALCDSKDCGKLSKD HFALAFHLITVQKLIKGIDPPLVLTPEKISPSNR ASLQKVTELTRKPVCIIFKGTILWRITDSIWMK HNRKRIWLRA
536	1886	A	4102	569	829	DHQK*KNIPCSWIGRINIVKMSILPKAIYRFSAI PIKIPMTFFTEI*S*NVYRTTKTQE*AKAILSKK EQNLEESHYLDFK*YYRAV
537	1887	A	4104	54	281	SIDCEHLIRRMLVLDPSKRLTIAQIKEHKWML IEVPVQRPVLYPQEQENEPSIGEFNEQVLRLM HSLGIDQQKTIE
538	1888	Α	4109	141	314	IRHIPLKIRSVVSHLKCFYKFILTFFFAGCSQPL VPRENITAWMNAIGLIITALPVS
539	1889	A	4111	268	1	ASRPWGHSYP*FNQQEVDTLKRPIASSEI*MM I*KFAT\KKSPGPYRFTAEFSHTFKEDLVPILW PLFPKIYREGTLPHSFYEASITL
540	1890	A	4142	198	2064	PEPGAGRAATPWGPLFWRGRGSGRCEKAAE AALGDFLGLHRRTQQPAVDRLLSDASAQWR VRGHGGVRESGRAPQQPGRRRGRRPRKRPR GRWRREGCGAGGRGVCVAAWSQRSIAGNN DYRLFHKMSNSHPLRPFTAVGEIDHVHILSEH IGALLIGEEYGDVTFVVEKKRFPAHRVILAAR CQYFRALLYGGMRESQPEAEIPLQDTTAEAFT MLLKYIYTGRATLTDEKEEVLLDFLSLAHKY GFPELEDSTSEYLCTILNIQNVCMTFDVASLY SLPKLTCMCCMFMDRNAQEVLSSEGFLSLSK TALLNIVLRDSFAAPEKDIFLALLNWCKHNSK ENHAEIMQAVRLPLMSLTELLNVVRPSGLLSP DAILDAIKVRSESRDMDLNYRGMLIPEENIAT MKYGAQVVKGELKSALLDGDTQNYDLDHG FSRHPIDDDCRSGIEIKLGQPSIINHVRILLWDR DSRSYSYFIEVSMDELDWVRVIDHSQYLCRS WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF ECMFTNKTFTLEKGLIVPMENVATIADCASVI EGVSRSRNALLNGDTKNYDWDSGYTCHQLG
541	1891	Α	4146	282	778	SGAIVVQLAQPYMIGSIRVLLWDCDDRSY GTLGYPNGARGQPQDNFFAHQ\VSHHPPISAC

NO. of NO. of Lot bod DI NO: beginning mulciotide totation operation of the control of the contr	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
muchocide coulde sequence USSN location corresponding to grow the corresponding to grow and the		, ,			1	1	
Docation Docation	nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine.
Seq. uence 99496 correspond g for first g cled residue credited peptide sequence credited	eotide	seq-	f	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine.
914		uence	1	09/496		to last amino	M=Methionine, N=Asparagine, P=Proline.
mino acid residue of sequence residue of	uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
Populide Sequence	1	1	1			of peptide	T=Threonine, V=Valine, W=Tryptophan.
	1					sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
HAESENRAFWODMKWRNFWGKSLEIVPVG TVNVSLPRROGHEFWKVTSCIANULSORW IEHYGEVLIRNTODSSCHCKITFCKAKYWSSN VIEWOGAVLSRSGRVLHRLFGKWHEGLYRG PIPGGOCIWKP IEHYGEVLIRNTODSSCHCKITFCKAKYWSSN VIEWOGAVLSRSGRVLHRLFGKWHEGLYRG PIPGGOCIWKP SVDAVVCNDIVESVETYTTILLEGALTHRIVA OPPROGOLISHLTCDSAPAGSGGTWSTSCR INHLIFRGGAQITHATFDDSFKAVLGDRILLT ANVSSEMITPRISKTTIPLEGALVAVTVV SH ITSYPQCLTQMYFLISFANVDTFLLPIMALDH YVAICSALQ*CSIITPSLCQGLPVLA*AGSSLIS PHYTVIMSRLAFCSSAQSHSFYDRAYLLMKIA CSHT*WNGHYFLGAVVLFLAPCALILVSYIRIA AALRIEPSTRRKACSISSHISLYTHAYLMKIA CSHT*WNGHYFLGAVVLFLAPCALILVSYIRIA AALRIEPSTRRKACSISSHISLYTHAYLMKIA CSHT*WNGHYFLGAVVLFLAPCALILVSYIRIA AALRIEPSTRRKACSISSHISLYTHAYLMKIA CSHT*WNGHYFLGAVVLFLAPCALILVSYIRIA AALRIEPSTRRKACSISSHISLYTHAY AND SSWDYRYSTPHANFFVEMFHHYVAGGLE LGGOLPTSTSHSAGITGVSHHAPPRLISSEGS LGGHLLCPMYFFLLCVVUSSSLAGEBAA LRVQKLWPAVVLSHLPVCWFLCSGIWSEVIE LVXGREGHLPWQGLAVCVVISSLAGEBAA LRVQKLWPAVVLSHLPVCWFLCSGIWSEVIE LVXGREGHLPWGAGAVSVFRIGAGSKSQGW ELELSGEPAPGWQULAGAYTYTQARYLRASE ANVGQPLRPVDR STPLOTE STPLO			İ	İ		Ì	
TWYSE_PREGIDITE_WIK_VISCIENVI_SQRW			<u> </u>	ļ	sequence		
IEHYGEVLIRNTQDSSCHCKITFCKAKYWSS VHEVQQAVLSRSGRVLHRIFGKAKYWSS VHEVQQAVLSRSGRVLHRIFGKAKYWSS VHEVQQAVLSRSGRVLHRIFGKAKYWSS SVDAVYCNDIVFSYRTTITLLEGA*LTHRYVA QDPKQGQLRSLHLTCDSAPAGSQGTWSTSC INHLIFRGGAQTFLATFDDSPKAVLGDRLLT ANNSSENITRTSKTTFQLESVKDAVYTVV SSI SSI 1893 A 4153 678 11			1	1		ł	HAESENFAFWQDMKWKNKFWGKSLEIVPVG
VHEVQQAVLSRSGRVLHRLFGKWHEGLYRG PTPGQCCIWEP PTPGQCCIWEP PTPGQCCIWEP PTPGQCCIWEP PTPGQCCIWEP PTPGQCCIWEP PTPGGCCIWEP PTPGGCCIWEP PTPGGCCIWEP PTPGGCCIWEP PTPGGCCIWEP PTPGGCCIWEP PTPGGCCIWEP PTPGGCCIWEP PTPGGCCIWEP PTPGGCCIPC PTPGCCIPC PTPGCCIP	İ		Ì				TVNVSLPRFGDHFEWNKVTSCIHNVLSGQRW
S42 1892 A 4147 44 433 SVDAYVCNDIVESYRTITILLEGA*LTHRYVA OPKGGQLRSLH.TCDSAPAGSGGTWSTSCK INHLIERGGA*QTFLATFDDSFKA*U-GRELLT ANN/SENNIPRTSKTTFQLELSVKDAVYTVV SH SVDAYCNDIVESYRDAVLORLLT ANN/SENNIPRTSKTTFQLELSVKDAVYTVV SH VYAICSALQ*CSIITPELCGGLPVLA*AGSSLL ANN/SENNIPRTSKTTFQLELSVKDAVYTVV SH VYAICSALQ*CSIITPELCGGLPVLA*AGSSLL PYHTVIMSRLARCSSAQISI*PYRDAYLLMKIA CHTVINGRIVACSSAQISI*PYRDAYLLMKIA CHTVINGRIVACSSAQISI*PYRDAYLLMKIA CHTVINGRIVACSSAQISI*PYRDAYLLMKIA CHTVINGRIVACSSCAQISI*PYRDAYLLMKIA CHTVINGRIVACSCSCAJISI*SVADYALLMKIA CHTVINGRIVACSCSCAJISI*SVADYALLMKIA CHTVINGRIVACSCSCAJISI*SVADYALLMKIA CHTVINGRIVACSCSCAJISI*SVADYALLMKIA CHTVINGRIVACSCAGISSHADAYLLMKIA CHTVINGRIVACSCAGISSHADAYLLMKIA CHTVINGRIVACSCAGISSAMISHVADYALLMXIA CHTVINGRIVACSCAGISSAGGISTAPARGE LICHALCLPMVPIFLACVALISSES SWDVRYSTPHANFVPMEPHIVAQAQILEL LOSGDLPTSTSISAGGTGVSSHAPPRI SSEGS LICHALCLPMVPIFLACVALISSES CHLCLELMVPI							IEHYGEVLIRNTQDSSCHCKITFCKAKYWSSN
1892 A	1						
SAS	542	1802	-	4147	44	422	
SHERRING AND SERNITE SKATTPQLESYKADAVYTVV SSH 1893 A 4153 678 11 TISYPQCLTQMYFLISPAN/DTFLLPIMALDH YVAICSALQ CSITIPELCQGLPVLA*AGSSLIS PYHTVIMSRLAFCSSAQISHFYRDAYLLMKIA CSHT*NQHYPLGAVVLELAPCALILVSYRIA AALRIPSTRRRACSICSSHLSLVTLFYQTV LGCPPPDSSAQDALATIMYVTVTSMLNPFIY SLMNKEVQEAVRRLFSRGSHSSWCW SLLYAQAGVQVINLSSLQQPGALQKSSHPSLP SSWDYRYSTPHPANFFYEMEFHHVAQAGLEL LGSGDLPTSTSHSAGTITOSHAPPLISSEGS LLGHLLCLPMVFPLLCVFVLISSSLAGEAAG LRVQKLWPAVVLSHLPVCWFHCSGIWSEVIE LKVGREGHVLPWQAHVVBF 545 1895 A 4160 1 412 HPLGLGLVFSEHSPQDKAADGSLLAPARGE DLEAGLKGSFMDGRLQASVSYFRIQRYGSAM QDTASAMPCLPYYTSHCFMAGGKSRSQGW ELELSGEPAFGWQVLAGYTYTAYLTASS SHOWN STRIPT STRIPT SHOPPORT SHOWN SHOW	342	1072	^	4147	44	433	
SSH							
SSH	}					·	
1893		ļ]				
YVAICSALQ*CSIITP/ELCQGLPVIA*AGSSLIS PVHTVIMSRLAFCSSAQISHFYRDAVLIMKIA CSHT*NQHVPLGAVVLFLAPCALILVSYTRIA AAILRIPSPTRRKACSICSSHLSLVTLFYGTV LGGCI*PPDSFSAQDALATMYVTSMLMPFIY SLMNKEVQEAVRRLFSRGSHSSWCW LICCLIPPOSFSAQDALATMYVTSMLMPFIY SLMNKEVQEAVRRLFSRGSHSSWCW LICCLIPPOSFSAQDALATMYVTSMLMPFIY SLMNKEVQEAVRRLFSRGSHSSWCW LICCLIPPOSFAQDALATMYVTSMLMPFIY SLMNKEVQEAVRRLFSRGSHSSWCW LICCLIPPOSFAQDALATMYSTLSGGS LLCHILLCLPMVPPLLCVPMEFHHVAQAGLEL LGSGDLPTSTSHSAGTIGVSHHAPPRLISSEGS LLCHILLCLPMVPPLLCVPMEFHHVAQAGLEL LGSGDLPTSTSHSAGTIGVSHHAPPRLISSEGS LLCHILLCLPMVPPLLCVPMEFGGWSEVIE LGSGDLPTSTSHSAGTIGVSHHAPPRLISSEGS LCHCHLCLPMVPPLCVPHCSGIWSEVIE LGSGDLPTSTSHSAGTIGVSHHAPPRLISSEGS LCHCHLCPMVPPLCVPHCSGIWSEVIE LGSGDLPTSTSHSAGTIGVSHHAPPRLISSEGS LCHCHLPMVPQLAVVLSGTYPQVSHAPAGGLAPAGE DLEAGILKGSFMOGH.QASVVFRIQRVGSAM QDTASAMPCLPYYPTSHCFMAGGKSRSQGW ELELSGEPAFGWVLAGYTYTQARYLRDASS ANVQQPLRYDR 546 1896 A 4174 1252 1190 FFQVFFILFLFFKTEFHSCCPGAVQWHDLDSL OPPPFRKFGSCLSLPSSWDYRHAPAHPANFV FLVPTOFHTVGQVASLELPTSGDTPASVASQSA GIGGVSHHAPPASQRCV GSHHASEGLAPOVPPAGGVSAQELTAPPQGG WGLGAPPAAPRESDERSAGNAVRSPSRQA RDSLQGRRLGGTRGAGPAGKGAGTMGPAS GFHSPPPRPHGPESPRSSCWHLLWFICPWPQ PSRLPRLTPAQULQGPGVAAPPCPP ALPKVLPPWTQKPLGCLQSPLPLLSPPPQIF ALPKVLPPWTQKPLGCLQSPLPLLSPPPQIF ALPKVLPPWTQKPLGCLQSPLPLLSPPPQIF ALPKVLPPWTQKPLGCLQSPLPLLSPPPQIF AASSLFWMGCQPPPHPGGLQSPPPLLSPPPQIF AASSLFWMGCQPPPHPGGLAVPPLICSAP PPPASGTSDSSDSSSPSASAARVWPAISPPP AARHRHPPEVTLSPCPFSCGFRLLGPRRQ ALUTGPHHPPSQQPQSSSWPHGFGGGPOPPL AASSLFWMGCQPPPHPGLWHRPLFUTGKA LASPLCFRAPGSLRCTPHPPPLFUTGKA ALUTGPHHPPSQPQPSSSSPPHGFGGFPPPLPALPP PPPRASTAGORYDPPSPLPCSSPPPLPCLSPPRQC LQPLHRAAQPLDFCCSLSPPGPPLPVTGKA LASPLCFRAPGSLRCTPHPSQCPPSSCWPHGFGGFPPPP PPPRAFSCGPFSSSOPPHHPPCLSAPPSRQC LQPLHRAAQPLDFCCSLSPPGPPLPPALPP PPPRASPASOPPYHAGPGSSSPLPGGTSP WPCPFHSPSPLPLSSPPPALRCPAQPFLHLP PPPPASPSCCPPLPLAGPGSSWPLGFTSP WPCPFHSPSPLPLSSPPPALRCPAQPFLHLP PPPPAPSPCCPLPPLAQPOSSPPPHARTLT PPPAPSPCCPLPPLAQPGSPPPHPPLPPLIPPGLPPAPSPLTLT PPPAPSPCCPLPPLAQPGSPPHPPLPFLLFFTLT PPPAPSPCCPLPPLAGPCSPPHLPPTLTCHA ASSLPTTATOTSTCHAMP PPPAPSPCCPLPPLAGPGSPSPPLARCPAPPTHLPP PPPPAPSPCCPLPPL	543	1893	A	4153	678	11	
PVHITVIMSRLAFCSSAQISHFYRDAYLLMKIA CSHT*NOHPVEGAVVLF1APCALIL\SYSTRIA AAILRIPSPTRRKACSICSSHLSLVTLFYGTV LGICI*PPDSFSAQDAIATIMYTVVTSMLNPFIY SALMKEVQEAVRRLFSRGSHSSWCW SLMKEVQEAVRRLFSRGSHSSWCW LLYAQAGVQ*LINLSSLQPQPAGILKQSSHPSLF SWDYRYSTPHANFFYEWBFHHVAQAGLEL LGSGDLPTSTSHSAGTIGVSHHAPPRLISSEGS LLGHLLCLPMVFPLLCVFVLISSLAGEAAG LRYQKLWPAVVLSHLPVCWFHCSGIWSEVIE LKVGREGHVLPWQAHVVEF PHEGIGLVSEBISPODKKAADGSILAPARGE DLEAGLKGSFMDGRLQASVSVRIQRYGSAM QDTASAMPCLPYYPTSHCFMAGGKSRSQGW ELELSGEPAPGWQVLAGYTYTQARYLRDASE ANVQQPLRPVDPR 546 1896 A 4174 1252 1190 FFQVFIFLETLFFFTEFHSCCFGAVQWHDLDSL QPPPPRFKGFSCLSLPSSWDYRHAPAHPANFY FLVETOFFLTWGQYASLELPTSGOTPASASQA GTGVSHHA*PRASGRCW 547 1897 A 4176 3029 1 AGFDGLAFASCQGARGQTRVPGAFSWLAP GSHASSEGLAPAAPRESDEKRAGSDAVRFSRGA RDSLGQRRLGGTRGAGPAGKGAQRTMGPAS GFHSFPRPHQEPSPRSSCWQHLLWTCPWPQ WGLAPPAAPRFESDEKRAGSDAVRFSRGA RDSLGQRRLGGTRGAGPAGKGAQRTMGPAS GFHSFPRPHQEPSPRSSCWQHLLWTCPWPQ PSRLPRLTPAQLLQGPGVLAAPPGP*HVPGFL AQSPWPLPSGRS**DPLHGGALVPLPQGGSP HTAPHCLPSVLSPAJQQPLLFTAGTSRSPPAS TMAPIPSALAVWEPAGSSPAJSAFADSAANVWPPGGSPPHVPGF- AQSPWPLPSGRS**PDLHGGALVPLPQGGSP HTAPHCLPSVLSPAJQQPLLFTAGTSRSPPAS TMAPIPSALAVWEPAGSSPASAPSAFAAANSFSRGA ALIVGPHHPP*SQQPQSQSVHPHGFGGPQPPL AASSLFWMCQPPPPHQGLWRRLPTVGKA LSSPLCFRFAPGSLRQTPLPQGSPPHLGSAP PPPASGTSDSSDSSSSSASAARVWPPASSPPP AASRLFWMCQPPPPHQFUWHRPLTGKA LSSPLCFRFAPGSLRQTPLPQGFPHLLGRPRPQ ALQTPRAWDLPPGSSPAPLCSGPELLGRPRRPQ ALQTPRAWDLPPGSSPAPLCSGPELLGRPRRPQ ALQTPRAWDLPPGSSPAPLCSGPELLGRPRRPQ ALQTPRAWDLPPGSSPAPLCSGPELLFPQCFPPLLGRPPSQC LQFLHRAAQPLDPCCSLSPFGPPLLGRPRPQ ALQTPRAWDLPPGSSPAPLCSGPELPPALPP PPPRSSGPHSPPLPLLSSPFPARRCPPPALPPALPP WFGPR*SPSPASSQPPYHAGLPGPGSSPLPPGL PQLPSLRSGSQQPLLFFQCPGPGAVWGRGSPQ PLSSPPPSPPASSGPTARYSVCL TPSRASSLPEVLASSLFRQSSGSUFLGFTSP MP*CFHRSSPPLPLLSSPFPALRCPAQPSPHLDF PPPAPSSGCPPLPPLAQPGPSPPLPPALPP PPPAPSPGCPLPPLAQPGPSPPLPALRPAQPFPHLHP PPPPAPSPGCPLPPLAQPGPSPPPARRCTLT PPPAPSPGCPLPPLAQPGPSPPPARRCTLT PPPAPSPGCPPLPPLAQPGPSPPPARRCTLT PPPAPSPGCPPLPPLAQPGPSPPPARRCTLT PPPAPSPGCPLPPLAQPGSSPPPARRCTLT		10,5	''	1133	076	**	VVAICSALO*CSITTP/ELCOCLEVILA*ACSSLIS
CSHT*NOHVFLGAVVLEILAPCALILVSYIRIA AARLRIPSPIRRKACSISLSUJVI.FYGTV LGICI*PPDSFSAQDAIATIMYTIVYTSMLNPFIY SLMNKEVQEAVRRLFSRGSHSSWCW LIVAQAGVQ*LNLSSLQFQPAGLKQSSHPSLP SSWDYRYSTPHPANFFVPMEFHHVAQAGLEL LGSGDLPTSTSHSAGITGVPUVPHCSGIWSEVIE LGSGDLPTSTSHSAGITGVPUVPHCSGIWSEVIE LKVGREGHVLPWQAHVVEF 545 1895 A 4160 1 412 HPLGIGLVFSEIPSPQDKKAADGSILAPARGE DLEAGLKGSSPMORLQASVSVFRIQRVGSAM QDTASAMPCLPYYPTSHCPMAGGKSRSQGW ELELSGEPAGWQVLAGYTYTQARYLRDASE ANVQQPLRPVDPR 546 1896 A 4174 1252 1190 FFQVFIFLFIFFKTEFHSCCPGAVQWHDDDSL QPPPRRKGPSCLSIPSSWDYRHAPAHPANFV FIVETOFILHVIQQUASLELPTSGDTFASVASQSA GTIGVSHHA*PRASGRRCG SHHASEGLAPVAPAGGVSAQELTAPPQEG WGLGAPPAAPREDEKRAGSDAVRSFSRGA RDSLQQRLGGTRQAGPAGKGAGRTMGPAS GFHSPPRPHGEPSRSSCWGHLWHCPWPQ PSRLPRLTPAQLLQGFGVLAAPPGFHVPGGL AGSPWLPPSGPSSPSPHUHGGALVPLPQGGSP HTAPHCLPSVLSPAIQQPLLPTAST/SSRSPPAS TMAPPSALAWEPAGSSPCWGHLWHCPWPQ PSRLPRLTPAQLLQGFGVLAAPPGFHVPGGL AGSPWLPPSGPSSPSPPHLIGGALVPLPQGGSP HTAPHCLPSVLSPAIQQPLLPTAST/SSRSPPAS TMAPPSALAWEPAGSSPCWHLWHCPWPQ PSRLPRLTPAQLLQGFGVLAAPPGFHVPGGL AGSSLFWMCQPPPPFPFQFFNGFGSPPPLPGGSAP/ PPPAGSTSDSSDSSSSSSASAARVWPPASSPPP AAPKPHPPSQPQPSSSVHPHGFGGPOPPL AASSLFWMCQPPPFPFQFFLUGRPRPQ ALVGPHHPPSQPCPSSSSPHLGGSAPAPA ALVGPHHPPSQPCPSSSSPHLGRRRPQ ALVGPHPPPSQPSPHSPTLAPPQAWPLELPPSRQC LQPLHRAAQPLDPCCSLSPFOPPLLGRRRPQ ALQTPRAWDLPPGSSPAARSPPPPL PPPRSPSQCPSSPPSRQCSPPSPRQCPSPPLCSRPCP PPPSRASSOPPHHGFGGPOSPLPPQL PPPSRPSSSOPPHSPTLAPPQAWPLELPPSRQC LQPLHRAAQPLDPCCSLSPFOPPLPPALPPQ PPPSRPSSSOPPLPSPRASSPPTARPCYPVYSVCL TPSRASSLPEVVLASSLPKIPQSSGSUCIGTTSP MP*CHRSSPPLPLSSPPPALRPCAPQPFLHLP PPPPPAPSPCCPPLALRPQAPSPPLP PPPPAPSPCCPLPPLAQPQPSPPPPHPPPPPPPPPPPPPPPPPPPPPPPPPPPPP							PVHTVIMSBI AECSSACISHEVDDAVI I MVIA
AAILRIPSPITRARACSICSSHLSLYTLFYGTV LGICI+PPDSFSAQDAIATIMTVVTSMINPFIY SLMNKEVQEAVRRLFSRGSHSSWCW LLYQAGGVPINLSSLQPQPAGEL(QSSHPSLP SWDYRYSTPIPANFPVEMEFHHVAQAGLEL LGSGDLPTSTSHSAGITGVSHHAPPRLISSEGS LLGHLLCLPMYFPLLCYNLSSLAGEAAG LRVQKLWPAVVLSHLPVCWFHCSGIWSEVIE LKVGREGHVLPWQAHVVEF 1895 A 4160 1 412 HPIGIGLVPSEIFSPQDKKAADGSILAPARGE DLEAGLKGSFMDGRLQASVSVFRQRVGSAM QDTASAMPCLPYYPTSIGFMAGGKSRSGGW ELELSGEPAPGWQVLAGYTYTQARYLRDASE ANVGQPLRPVDPR 546 1896 A 4174 1252 1190 FFQVFIFLFLIFFKTEFHSCCPGAVQWHDLDSL QPPPPRFKGFSCLSLPSSWDYRHAPAHPANFV FLVETGFILTHVGQVASLELPTSGDTPASVASQS GITGVSHHA*PRASGRRCW 547 1897 A 4176 3029 1 AGPDGLAAPASCGGARGQTRVPGAFSWLAP GSHHASEGLAPGVPPAGGVSAQELTAPPQEG WGLGAPPAAPRPESDEKRAGSDAVRSFSSRA ADSLGGRRLGGTRGAGPAGKGAQRTMGPAS GFHSFPRPHQEPSPRSCWQHLLWHCPWPQ PSLIPRLTPAQLLQCPQAPAPGVSAQELTAPPQEG WGLGAPPAAPRPESDEKRAGSDAVRSFSSRA ADSLGGRRLGGTRGAGPAGKGAQRTMGPAS GFHSFPRPHQEPSPRSCWQHLLWHCPWPQ PSLPRLTPAQLLQCPQPPHPQGSP HTAPHCLPSVLSPAIQQPLLPTASTSSRSPPAS TMAPIPSALAWEPAGSPQLSAPADASSPLP ALFKVLPPWTQKPLGCLQSSPLDLSSPPOU RCPPACSPAAASSFSFESQPCSPAPSKASPAPA ALIVGPHIPP*SQQPQSQSVHPHGPGGGPQPPL AASSLFWMCQPPPHPQFHIPRGLSAP PPPASGTSDSSDSRSPSASARARWPPAISPPP AARRPHPPFYFLSCGFFRLLGRRRPQ ALGYPRAWDL.PPGGSPAPLAGSPPEDLPAPPLP PPPAPSFASGQPYHAGLFGGPGSSPLPCSQP AARRPHPPFYFLSCGFFRLLGRRRPQ ALGYPRAWDL.PPGGSPAPLCSGPELP+APPPLP PPPAPSPASGQPPJHAGLFGGPGSSPLCPTSP WPGRP*SPASSGQPVHAGLFGGPGSSPLCPTSP WPGRP*SPASSGQPVHAGLFGGPGSSPLCPTSP WPGRP*SPASSSCPPVHAGLFGFQSSSPLCPTSP WPGPR*SPASSSCPPVHAGLFGFQSSSPLCPTSP WPGPR*SPASSSPPVHAGLFGFQSSSPLCPTSP WPGPR*SPASSSPPVHAGLFGFQSSSPLCPTSP MPCCHRPSPPLATGTTFPVARSSLSPGTAPYSVCL TPSRASSSLPEVVLASSLSPGTAPSSCLL TPSRASSSLPEVVLASSLSPGTAPSSVCL TPSRASSSLPEVVLASSLSPGTAPSSVCL TPSRASSSLPEVVLASSLSPGTAPSSVCL TPSRASSSLPEVVLASSLSPGTAPSSVCL TPSRASSSLPEVVLASSLSPGTAPSSVCL TPSRASSSLPEVVLASSLSPGTAPSSVCL TPSRASSSLPEVVLASSLSPGTAPSSVCL TPSRASSSLPEVVLASSLSPGTAPSSVCL TPSRASSSLPEVVLASSLSPGTAPSSVCL TPSRASSSLPEVALASSLSPGTAPSSVCL TPSRASSSLPEVALASSLSPGTAPSSVCL TPSRASSSLPEVALASSLSPGTAPSSVCL TPSRASSSLPEVALASSLSPGTAPSSVCL		j.	J		٠	i	CSHT*\NOHVELGAVVLELAPCALILYSVIPLA
LGICI*PPDSFSAQDAIATIM/TVVTSMI.NPFIY SLMNKEVQEAVRELPSCSHSSWCW 1894 A 4158 3 538 LLYAQAGVQ*INLSSLQPQPAGLKQSSHPSLP SSWDYRYSTPHPANFFVEMEFHHVAQAGLEL LGSGDLPSTSSHSAGTGVSHHAPPRLISSEGS LLGHLLCLPMYFPLCVFVLISSSLAGEEAAG LRVQKLWPAVVLSHLPVCWFHCSGIWSEVIE LKVGREGHVLPWQAHVVEF 1895 A 4160 1 412 HPLGLGLVFSEITSPQDKKAADSSILAPARGE DLEAGLKGSFMDGRLQASVSVFRIQRVGSAM QDTASAMPCLPYYPTSHCFMAGGKSRSQGW ELELSGEPAPGWQVLAGYTYTCARYLRDASE NVQQPLRPVDPR 546 1896 A 4174 1252 1190 FFQVFFLFLFFKTEFHSCCPGAVQWHDLDSL QPPPPRKFGFGSLSLSPSWDYRHAPHPANFV PLVETGFLHTWQDASILEDFISQDTPASIASQSA GITGVSHHA*PRASGRCW QSHAPASCGGARGQTRVPGAFSWLAP GSHHASEGLAPGVPPAGGVSAQELTAPPQEG WGLGAPPAAPPESDEKAGSDAVRSSRGA RDSLGQRRLGGTRGAGPAGKGAQRTMGPAS GFHSFPRPHGPERSPSSCWLHWHCPWPG PSRLPRLTPAQLLQQPGVLAAPPCP*HVPGFL AQSPWPLPSGPRSPD*DLHQGALVPLPQGGSP HTAPHCLPSVLSPALQPCPLPATSTSSRSPAS TMAPPSALAVWEPAGSSPQLSAPADSSVLP ALPKVLPPWTQRFLGCLQSPLPLSPFDQU RCPPACSPAAASSSFSPCGPPSAPSKASPAPA ALIVGPHHPP*SQQPQSQSVIHPHGGGGPPPL AASSLFWMFCQPPPPHPQFLWHRPLPVTGKA LASSPLCFRPAPGSLRQTPLPPGFHPPGLSAP/ PPPASGTSDSSDSRSPSASAARWPPASPPPP AARRHPPEYFLSCFGFFRLLGRFRRPQ ALQTPRAWDLPPGGSSPAPLCSGPELP*APPPLP PPPRVSH_CSGGPPSAPSKASPAPA ALIVGPHHPP*SQQPQSQSVIHPHGGGGPPUP PPPRVSH_CSGGPPSAPSKASPAPA ALIVGPHHPP*SQQPQSSVIPPGLPVPALPS PPPASGTSDSSDSRSPASAARWPPASPPPP AARRHPPEYFLSCFGFFRLLGRFRRPQ ALQTPRAWDLPPGGSSPAPLCSGPELP*APPPLP PPPRVSH_CSGGPPSAPSKASSPAPA ALIVGPHHPP*SQQPQSQSVIPPGGPPVPALPS PPPASGTSDSSDSRSPASASRSVPAPPPLPPPLPPPPLPPPPARPSPLSSGPPPLPVPALPS PPPRVSH_CSGGPPPARTSCGPPLPPARPSPPPLPPPLPPPPPPAPPLPPPPPARPSPLSSGPPPLPVPALPS PPPRVSH_SGGPPSPASSSQPLCFTSPQC LQPLHLRAAQPLDPCCSLSPPGPPLPVPALPS PPPRPASGGPPPARSSSSPAPSSSAPSSSPAPSVCLL TPSRASSLPEVVLASSLPKIPQSSGSPLGPTSP MPCCHRFSPPLASSPPPARSSLPFOTAPSVCLL TPSRASSLPEVVLASSLPKIPQSSGSPLGPTSP MPCCHRFSPPLASSPFPALRSTLT PPPPAPSPPSCOPLPPARSSLSSPGTAPSVCLL TPSRASSLPEVVLASSLPKIPQSSGSPLGPTSP MPCCHRFSPPLASSPFPALRSTLT		1					AAII RIPSPTRRRK ACSICSSHI SI VTI EVCTV
SLMNEVQEAVRILESRGSHSSWCW							LGICI*PPDSFSAODAIATIMYTVVTSMI NPFIV
S44			İ				
SSWDYRYSTPHPANFFVEMEFHHVAQAGLEL LOSGLDTTSTHSAGTGVSHHAPPRLISSEGS LLGHLLCLPMVFPLCVFVLISSSLAGEEAAG LRVQKLWPAVVLSHLPVCWFHCSGIWSEVIE LKVGREGHVLPWQAHVVEF 545 1895 A 4160 1 412 HPLGIGLVPSEIFSPQDKKAADGSILAPARGE DLEAGLKOSFMORLQASVSVFRIQRVGSAM QDTASAMPCLPYYPTSHCFMAGGKSRSQGW ELELSGEPAFGWQVLAGYTYTQARYLRDASE ANVGQPLRPVDPR 546 1896 A 4174 1252 1190 FFQVFIFLFIFFKTEFHSCCPGAVQWHDLDSL QPPPPRRKGFSCLSLPSSWDYPHAPAHPANFV FLVEFGFLHVVGQVSASLELPTSGDTPASASQSA GITGVSHHA*PRASGRRCW 547 1897 A 4176 3029 1 AGPDGLAAPASCQGARGQTRVPGAFSWLAP GSHHASEGLAPCVPPAGGVSAQELTAPPQEG WGLGAPPAAPRESDEKRAGSDAVRSFSRGA RDSLGQRRLGGTRGAGPAGKGAQRTMGPAS GFHSFPRPHQEFSPRSSCWQHLLWHCPWPQ PSRLPRLTFAQLLQGFQVLAAPFGP*HVPGFL AQSPWPLSGPRS*PDHLGALAVPLPQGGSP HTAPHCLPSVLSPAIQQPULPTAST/SSRSPPAS TMAPIPSALAVWEPAGSSPQLSAPADSSVLP ALPKYLPWTGKALAVWEPAGSSPQLSAPADSSVLP ALPKYLPWTGKPLLGCLCQSSIPLLSPPDQV RCPPACSPAAASSFSESQPCPSAPSKASPAPA ALIVQPHHPP*SQQPQSQVHPHGPGGPQPPL AASSLFWMFCQPPPPHPQFILWHRPLPVTGKA LASPLCRPAPAGSLRQTILPSPPOHIPRGGLSAP/ PPPAGTSDSSDSSTSSASAARVWPPASPPPP AARRHPPPEYFLSPCPFSCGFPRLLGRPRRPQ ALQTFRAWDLPPGSSPALLCSGFELP*APPPLP PPPASPALCSGPHSPSQVVCGWFRCV*GHPIP RPVGHS*SGPPHSPL*APPQAWPLELPPSRQC LQPLHRAAQPLDPCCSLPPSRQCL LQPLHRAAQPLDPCCSLPPSRQCL LQPLHRAAQPLDPCCSLSPPGPPLPVALPS WPGRP*SPSPASSQPYHAGLDGPQSSPLPGL PQLPSLRSGSQQPLLFFCCPGPGAV WGKGSPQ PLSPHPPPPARTCTPTPVASRSLSPCTAPYSVCL TPSRSASSLPEVVLASSLPKIPQSSGSPLEGTSP MP*CFHRSPSPPLLSSPPPAURRQAPQPTPLHLP PPPAPASPGCCLLPPLAQQHQESPSPSHARSTLT	544	1894	Α	4158	3	538	
LGSGDLPTSTSHSAGITGVSUSHHAPPRLISSEGS LLGHLLCLPMVFPLLCVFVLISSSLAGEAAG LRVQKLWPAVVLSHLPVCWPHCSGIWSEVIE LKVGREGHVLPWQAHVVEF 545 1895 A 4160 1 412 HPLGIGLVPSEIFSPQDKKAADGSILAPARGE DILEAGLKGSFMDGRLQASVSYFRIQRVGSAM QDTASAMPCLPYYPTSHCFMAGGKSRSQGW ELELSGEPAPGWVLAGYTYTQARYLRDASE ANVGQPLRPVDPR 546 1896 A 4174 1252 1190 FFQVFIELTLFFEKTEFHSCCPGAVQWHDLDSL QPPPPRFKGFSCLSLPSSWDYRHAPAHPANFV FLVETGFLITVGQVASSLELPTSGDTPASSASQSA GTGVSSHAPPRASGRRCW 3029 1 AGPDGLAAPASCQGARGQTRVPGAFSWLAP GSHHASEGLAPGVPPAGGVSAQELTAPPQEG WGLGAPPAAPRESDEKRAGSDAVRSFSRGA RDSLQGRRLGGTRGAGPAGKGAQRTMGPAS GFHSFPRPHQEFSPSESKWHLLWHCPWPQ PSRLPRLTPAQLLQGPGVLAAPPGPHVPGFL AQSPWPLPSGPRSP*DPLHQGALVPLPQGGSP HTAPHCLPSVLSPAQQPLLTATSISSRSPPAS TMAPPSALAVWEPAGSSPQLSSAPADSSUP,P ALPKVLPPWTOKRLUCQSPPLLJSPPOQI RCPPACSPAAASSFSFESQPCPSAPSKASPAPA ALIVGPHHPP*SQQPQSQSVPHPHGPGGPQPPL AASSLFWMFCQPPPHPQFI WHRPLPVTGKA LASVLCRPAPGSSAARVWPPASSPPP AARRPHPPPFYFLSPCPFSCGFPRLLGRPRRQ ALQTPRAWDLPPGSSPAPLCSGFELP*APPPLP PPPASGTSDSSDSSPSASAARVWPPASSPPP AARRPHPPPSYPLSPCPPSQCFPLLGRPRRQ ALQTPRAWDLPPGSSPAPLCSGFELP*APPPLP PPPASGTSDSSDSPSPSASAARVWPPASSPPL PPRVGHS*SGPPHSPPL*APPQAWPLELPPSRQC LQPHLRAAQPLDPCCSLSGFELP*APPPLP PPPASGTSDSSDSPSPSAQPPLLGRPRRQC LQPLHLRAAQPLDPCCSLSGFELP*APPPLP PPPASGTSDSSDSPSPSAQPPLLGRPRRQC LQPLHRAAQPLDPCCSLSGFELP*APPPLP PPSVGHS*SGPPHSPPL*APPQAWPLELPPSRQC LQPLHRAAQPLDPCCSLSGFELP*APPPLP PPSVGHS*SGPPHSPPL*APPQAWGLSPPSPPLPLSSPPPAURPQAPGPPLLLP PPTSRAASSLPEVVLASSLPKIPQSSGPLPGTSP WPGFR*SPSPASSQPPLYAGSLPGTSPSVCL TPSRSASSLPEVVLASSLPKIPQSSGSPLPGTSP WPGFRF*SPSPASSLPCCLPPLARSSTLPGTSPSVCL TPSRSASSLPEVVLASSLPKIPQSSGSPLPGTSP WPGFRF*SPSPASSLPFAURPQAPGFPLHLP PPPAPSFCCLPPLAGQPGPSPSPAMRSTLT	1	ĺ	1	1 1			SSWDYRYSTPHPANFFVFMEFHHVAOAGLEL
LLGHLLCLPMYFPLLCVFVLISSSLAGERAGG LRVQKLWPAVYLDLPVCWFHCSGIWSEVIE 545 1895 A 4160 1 412 HPLGLGLVPSEIFSPORKADGSILAPARGE DLEAGLKGSFMDGRLQASVSYPRIQRVGSAM QDTASAMPCLPYYPTSHCFMAGGKSRSQGW ELELSGEPAPGWQVLAGYTYTQARYLRDASE ANVQPLRPVDPR 546 1896 A 4174 1252 1190 FFQVFIFLIFFKTEFHSCCPGAVQWHDLDSL QPPPPRFKGFSCLSLPSWDYRHAPAHPANFV PLVETGFLHVUQVASLELPISODIPASIASQSA GITGYSHHA*PRASGRCW 547 1897 A 4176 3029 1 AGPDGLAAPASCQGARGQTRVPGAFSWLAP GSHHASEGLAPGVPPAGGVSAQELTAPPQEG WGLGAPPAAPRPESDERRAGSDAVRSFSRGA RDSLGQRRLGGTRGAGPAGKGAQRTMGPAS GFHSFPPRPQEFSPRSSCWQHLLWHCPWPQ PSRLPRLTPAQLLQCPGVLAPPGP*HVPGFL AQSPWPLPSGFRSPDPLHQGALVPLPQGGSP HTAPHCLPSVLSPAIQQPLLFTAST/SSRSPPAS TMAPIPSALAVWEPAGSSPQLSSAPADSSVIP ALPKVLPPWTQKPLLGCLQSPLPLLSPPDQI/ RCPPACSPAAASSFSFESOPCPSAPSKASPAPA ALJVVGPHHPP*SQQPQSOSVHPHIPGGGOPPL AASSLFWMFCQPPPPHPQFLWHRPLPVTGKA LASVPLCFRRAPGSLRQTIPPQFHIPRFGLSAP/ PPASGTSDSSSDSRSPSASAARVWPPAISPPPP AARHRPHPEYFLSPCFPSCGFPRLLGRPRRQ ALQTPRAWDLPPGSSPAPILCSGPELP*APPPLP PFRVA*LGSGHPPPSAQVPGLW*RCV*GHPIP RPVGHS*SGPPHSPPL*APPQAWPLELPPSRQC LQPLHRAAQPLOCSCLSPPPGPPLPVALPS WPGRP*SPSPASSQPPYHAGLPGPQSSPLPPGL PQLPSLRSGSQQPLLFFQCOPGAVWGKGSPQ PLSFMPPP/ARTOTTPVASRSLSPGTAPYSVCL TPSRASSLFEVVLASSLPKIPQSSGSVLGPTSP MP*CFHRPSPPLPLAQSHQHQPSPSPANSTLT	1		1				
LRVQKLWPAVVLSHLPVCWFHCSGIWSEVIE LKVGREGHVLPWQAHVVEF 1895 A 4160 1 412 HPLGI.GLVPSEIFSPQDKKAADGSILAPARGE DLEAGLKGSFMDGRL.QASVSVFRIQRVGSAM QDTASAMPCLPYYPTSHCFMAGGKSRSQGW ELELSGFPAPGWQVLAGYTYTQARYLRDASE ANVGQPLRPVDPR 546 1896 A 4174 1252 1190 FPQVFIFLFLIFFKTEFHSCCPGAVQWHDLDSL QPPPRFKGFSCLSLPSSWDYRHAPAHPANFV PLVETGFLHVIQVASLELPTSGDTPASIASQSA GITGYSHHA*PRASGRRCW GITGYSHHA*PRASGRRCW GIGAPPAARPESDEKRAGSDAVRSFSRGA RDSLGQRRLGGTRGAGPAGKGAQRTMGPAS GFHSPPRPHQEPSPRSSCWGHLLWHCPWPQ PSRLPRALDLQGPGVLAAPPESPHVFGIFL AQSPWPLPSGPRSP*DPLHGGALVPLPQGGSP HTAPHCLPSVLSPAIQQPLLPTAST/SSRSPPAS TMAPIPSALAVWEPAGSSPQLSSAPADSSVILP ALIVGPHHPP*SQQPQSQSVHPHGPGGPQPPL AASSLFWMFCQPPPHIPPGILWHRPLPVTGKA LASSLFWMFCQPPPHIPPGILWHRPLPVTGKA LASSLFWMFCQPPPHIPPGILWHRPLPVTGKA LASSLFWMFCQPPPHIPPGILAPP AARRRHPPPSTSCFPRLLGRPRRPQ ALIVGPRAWDLPGGSSPALSGFPSAPLASGFELP*APPPL PPPASGTSDSSDRSSPSASAARVWPPALSP PPPARSTDSSSPSAPLCSGFPLP*APPPLP PPPRVA*LGSGHPPSAQVPGLW*RCV*GHPIP RPVGHS*SGPPHSPPL*APPQAWPLELPPSRQC LQPLHRAAQPLCSLSSPFOPLYVBLPS WPGRP*SPSPASSQPPYHAGLPGPQSSPLPPGL PQLPSLRSGSQQPLLFFCCPGPGAVWGKGSPQ PLSFHPPP/ARTQTTPPVASRSLSFGTAPYSVCL TPSRASSLPEVVLASSLPKIPQSSGSVLCPTSP MP*CFHRPSPPLALSSIPPFSALRPQAPQFPHLIPP P*PPAPSPCCPLPLAQQHQPSPSPSPHARSTLT							
LKVGREGHVLPWQAHVVEF		ļ					
DLEAGLKOSEMOGRLQASVSVFRIQRVOSAM QDTASAMPCLPYYPTSHCFMAGGKSRSQGW ELELSGEPAPGWQVLAGYTYTQARYLRDASE ANVQPLRPVPDPR 546 1896 A 4174 1252 1190 FFQVFIFLFLIFFKTEFHSCCPGAVQWHDLDSL QPPPFKGFSCLSLPSSWDYRHAPAHPANFV FLVETGFLHVGQVASLELPTSGDTPASVASQSA GTTGVSHHA*PRASGRRCW 547 1897 A 4176 3029 I AGPGLAAPASCQGARGQTRVPGAFSWLAP GSHHASEGLAPGVPPAGGVSAQELTAPPQEG WGLGAPPAAPRESDEKRAGSDAVRSFSRGA RDSLGQRRLGGTRGAGPAGKGAQRTMGPAS GFHSFPPRPHQEPSPRSSCWQHLLWHCPWPQ PSRLPRLTPAQLLQGPGVLAAPPGF*HVPGFL AQSPWPLPSGPRSSP*DPLHQGALVPLPQGGSP HTAPHCLPSVLSPAIQQPLLPTAST/SSRSPPAS TMAPIPSALAVWEPAGSSPQLSAAPADSSIPLP ALPKVLPPWTQKPLLGCLCQSPLPLLSPPDQI/ RCPPACSPAASSFSFESQPCPSAPSKASPAPA ALUVGPHHPP*SQQPQSQSVHPHGPGGPQPPL AASSLFWMFCQPPPPHPQFLWHRPLPVTGKA LASPLCFRPAPGSLRQTPLPPQFHIPRPGLSAP/ PPPAGGTSDSSDSRSPSASAARVWPPAISPPP AARHRPHPPFYFLSPCPFSCGFPRLLGRPRRPQ ALQTRRAWDLPPGSSPAPLCSGPELP*APPPLP PPPRVA*LGSGHPPSAQVPGLW*RCV*GHPIP RPVGH*S*GPPPBSPALSQPQCV*GHPIP RPVGH*S*GPPPBPPLP*APPQAWPLELPPSRQC LQPLHLRAAQPLDPCCSLSPPGPPLPVALPS WPGRP*SPSPASSQPPYHAGLPGPSSPLPGL PQLPSLRSGSQQPLLFFQCPGPGAVWGKGSPQ PLSPHPPPFARTQTFPVASRSLSPGTAPYSVCL TPSRSASSLPEVVLASSLJPKIPQSSGSPLGFTSP MP*CFHRPSPPLP/LSSPFPAURRQAPQFPLHLP P*PPAPSPSCCPPPLAQQHQPSPPSPPARSTLT	L						LKVGREGHVLPWQAHVVEF
DLEAGLKGSFMOGRLQASVSVFRIQRVGSAM QDTASAMPCLPYYPTSHCFMAGGKSRSQGW ELELSGEPAFGWQVLAGYTYTQARYLRDASE ANVQPLRPVDPR 546 1896 A 4174 1252 1190 FFQVFIFLFIFFKTEFHSCCPGAVQWHDLDSL QPPPPRFKGFSCLSLPSSWDYRHAPANFVV FIVETGFLHVGQVASLELPTSGDTPASASQSA GITGVSHHA*PRASGGRCW 547 1897 A 4176 3029 1 ACPDGLAAPAGCGARGGTRVPGAFSWLAP GSHHA*SEGLAPGVPPAGGVSAQELTAPPQEG WGLGAPPAAPRESDEKRAGSDAVRSFSRGA RDSLGQRLIGGTRGAGPAGKGAQRTMOPAS GFHSFPPRPHQEPSPRSSCWQHLLWHCPWPQ PSRLPRLTPAQLLQGPGVLAAPPGF*HVPGFL AQSPWPLPSGPS*PDFLPGGAVPPLGGGSP HTAPHCLPSVLSPAIQQPLLPTAST/SSRSPAS TMAPIPSALAVWEPAGSSPQLSAPADSSIVLP ALPKVLPPWTQKPLLGCLGQSPLPLLSPPDQW RCPPACSPAAASSFSFESQPCPSAPSKASPAPA ALVIGPHHPP*SQQPQSQSVHPHGPGGPQPPL AASSLFWMFCQPPPPHPQFLWHRPLPVTGKA LASSPLCFRPAPGSLRQTPLPPQFHIPRPGLSAP/ PPPASGTSDSSDSRSSPASAARVWPPAISPPP AARRPHPPSFVLSPCFFSGFPRLLGRPRRPQ ALQTPRAWDLPPGSSPAPLCSGPELP*APPPLP PPPRVA*LGSGHPPSAQVPGLW*RCCV*GHPIP RPVGHS*SGPPHSPPL*APPQAWPLELPPSRQC LQPLHLRAAQPLDPCCSLSPPGPPLPURSPPP PFRVA*LGSGHPPSAQVPGLW*RCCV*GHPIP RPVGHS*SGPPHSPPL*APPQAWPLELPPSRQC LQPLHLRAAQPLDPCCSLSPPGPPLPVALPS WPGRP*SPSPASSQPPYHAGLPGPGSSPLPGTL PSRSASSLPEVVLASSLPKIPQSSGSVLGFTSP MP*CFHRPSPPLP/LSSPFPAVLRQAPQFPLHLP P*PPAPSPGCPLPPLAQQHQPSPPSPPHARSTLT	545	1895	Α	4160	1	412	HPLGLGLVPSEIFSPQDKKAADGSILAPARGE
QDTASAMPCLPYYPTSHCFMAGGKSRSQGW ELELSGEPAPGWQVLAGYTYTQARYLRDASE ANVQQPLRPVDPR 546 1896 A 4174 1252 1190 FFQVFIFLFIFFKTEFHSCCPGAVQWHDLDSL QPPPPRKGFSCLSLPSSWDYRHAPAHPANFV FLVETGFLHVGQVASLELPTSGDTPASVASQSA GITGVSHHA*PRASGRCW 547 1897 A 4176 3029 1 AGPDGLAAPASCQGARGQTRVPGAFSWLAP GSHHASEGLAPGVPPAGGVSAQELTAPPQEG WGLGAPPAAPRPESDEKRAGSDAVRSFSRGA RDSLGQRRLGGTRGAGPAGKGAQRTMGPAS GFHSFPRPHQEPSPRSSCWQHLLWHCPWPQ PSRLPRLTPAQLLQGPGVLAAPPGF*HVPGFL AQSPWPLPSGPRSP*DPLHQGALVPLPQGGSP HTAPHCLPSVLSPAIQQPLLPTAST/SSRSPPAS TMAPPSALAVWEPAGSSPQLSSAPADSSVPLP ALPKVLPPWTQKPLLGCLCQSPLPLLSPPDQI/ RCPPACSPAAASSFSESQCPSAPSKASPAPA ALIVGPHHEPP*SQQPQSQSVHPHGPGGPQPPL AASSLFWMFCQPPPHPQFLWHRPLPVTGKA LASSLFWMFCQPPPHPQFLWHRPLPVTGKA LASSLFWMFCQPPPHPQFLWHRPLPVTGKA LASSLFWRFQPPSPSSASAARVWPA\SPPPP AARHRPHPEYFLSPCPFSCGFPRLLGRPRRPQ ALQTPRAWDLPGSSPAPLCSGPELP*APPPLP PPPPASGTSDSSDSSPSPAQWFLGW*RCV*GHPIP RPVGHS*GSPPHSPPL*APPQAWPLELPPSRQC LQPLHRAAQPLDPCCSLSPPGPPLPVPALPS WPGRP*SPSPASSQPYHAGLGGPQSSPLPPGL PQLPSLRSGQQPLLFFQCPGFGAVWGKGSPQ PLSPHPPPP/ARTQTFPVASRSLSPGTAPYSVCL TPSRSASSLPEVVLASSLPKIPQSSGSWLGPTTSP MP*CFHRSPPLASSPPPALRPQAPQFPLHLP P*PPAAPSFGCPLPPLAQQHQPSPPSPHARSTLT							DLEAGLKGSFMDGRLQASVSVFRIQRVGSAM
546 1896 A 4174 1252 1190 FFQVFIFLFLIFFKTEFHSCCPGAVQWHDLDSL QPPPPRFKGFSCLSLPSSWDYRHAPAHPANFV FLVETGFLHVIGQVASLELPTSGDTPAS/ASQSA GITGVSHHA*PRASGRRCW 547 1897 A 4176 3029 1 AGPDGLAAPASCQGARGQTRVPGAFSWLAP GSHHASEGLAPGVPPAGGVSAQELTAPPQEG WGLGAPPAAPRPESDEKRAGSDAVRSFSRGA RDSLGQRRLGGTRGAGPAGKGAQRTMGPAS GFHSFPPPHQEFSPRSSCWQHLLWHCPWPQ PSRLPRLTPAQLLQGPGVLAAPPGP*HVPGFL AQSPWPLPSGPRS*PDLHQGALVPLPQGGSP HTAPHCLPSVLSPAIQQPLLPTAST/SSRSPPAS TMAPIPSALAVWEPAGSSPQLSSAPADSSIPLP ALPKVLPPWTQKPLLGCLCQSPLPLLSPPDQI/ RCPPACSPAAASSFSFESQPCPSAPSKASPAPA ALIVIGPHHPP*SQQPQSQSVHPHGPGGPQPPL AASSLFWMFCQPPPPHPQFLWHRPLPVTGKA LASSLFWMFCQPPPPHPQFLWHRPLPVTGKA LASSLFCFRRAPGSLRQTTPLPPGFHIPRFGLSAP/ PPPASSDSDSRSPSASAARVWPA\SPPPP AARHRPHPPEYFLSPCPFSCGFPRLLGRPRRPQ ALQTPRAWDLPPGSSPAPLCSGPELP*APPLP PFPVA*LSGGHPPSAQVPGLW*RCV*GHPIP RPVGHS*SGPPHSPL*APPQAWPLELPPSRQC LQPLHLRAAQPLDPCCSLSPPGPPLPVALPS WPGRP*SPSPASSQPYHAGLFGPQSSPLPGG PQLPSLRSGSQQPLLFFQCCPGPGAVWGKGSPQ PLSPHPPPP/ARTQTFPVASRSLSPGTAPYSVCL TPSRSASSLPEVVLASSLPKLPQSSGSIPLGFTSP MP*CFHRSPSIP/LSSPFPALRPQAPGFPLHLP PPPAPSPGCPLPPLAQQHQPSSPSHARSTLT							QDTASAMPCLPYYPTSHCFMAGGKSRSQGW
546 1896 A 4174 1252 1190 FFQVFIFLIFFKTEFHSCCPGAVQWHDLDSL QPPPRRKGFSCLSLPSSWDYRHAPAHPANFV FLVETGFLHVVGQVASLELPTSGDTPAS\ASQSA GITGVSHHA*PRASGRCW 547 1897 A 4176 3029 I AGPDGLAAPASCQGARGQTRVPGAFSWLAP GSHHASEGLAPGVPPAGGVSAQELTAPPQEG WGLGAPPAAPRESDEKRAGSDAVRSFSRGA RDSLGQRRLGGTRGAGPAGKGAQRTMGPAS GFHSFPPRPHQEFSPRSSCWQHLLWHCPWPQ PSRLPRLTPAQLLQGPGVLAAPPGP*HVPGFL AQSPWPLPSGPRSP*DPLHQGALVPLPQGGSP HTAPHCLPSVLSPAIQQPLLPTAST/SSRSPPAS TMAPIPSALAVWEPAGSSPQLSSAPADSSIPLP ALPKVLPPWTQKPLLGCLCQSPLPLLSPPDQI/ RCPPACSPAASSFSFESQPCPSAPSKASPAPA AL\IVGPHHPP*SQQPQSQSVHPHGPGGPQPPL AASSLFWMFCQPPPHPQFILWRPLPVTGKA LASPLCFRAPGSLRQTPLPPQFHIPRPGLSAP/ PPPASCTSDSSDSSDSSPSASAARVWPPA\SPPPP AARHRPHPEYFLSPCPFSCGFPRLLGRPRRPQ ALQTPRAWDLPPGSSPAPLCSGFELP*APPPLP PFPRVA*1GSGHPPSAQVPGLW*RCV*GHPIP RPVGHS*SGPPPSPPAARSPAQVPGLW*RCV*GHPIP RPVGHS*SGPPPSPASQPPYHAGLPGPQSSPLPGL PQLPSLRSGSQQPLLFFQCPGPGAVWGKGSPQ PLSPHPPPP/ARTQTFPVASRSLSPGTAPYSVCL TPSRSASSLPEVVASSLPKIPQSSGSPLEIPTSP MP*CFHRPSPPLP/LSSPFPA\LRPQAPFPHLP P*PPARSPGCPLPPLAQQHQPSPPSPHARSTLT			Į			:	ELELSGEPAPGWQVLAGYTYTQARYLRDASE
OPPPREKGFSCLSLPSSWDYRHAPARIPANFV FLVETCFLHYUGQXASLELPTSCDTPAS\ASQSA GITGVSHHA*PRASGRCW 547 1897 A 4176 3029 1 AGPGLAAPASCQGARGQTRVFGAFSWLAP GSHHASEGLAPGVPPAGGVSAQELTAPPQEG WGLGAPPAAPRPESDEKRAGSDAVRSFSRGA RDSLGQRRLGGTRGAGPAGKGAQRTMGPAS GFHSFPRPHQESPRSSCWQHLLWHCPWPQ PSRLPRLTPAQLLQGPGVLAAPPGP*HVPGFL AQSPWPLPSGPRSP*OPLHQGALVPLPQGGSP HTAPHCLPSVLSPAIQPLLPTAST/SSRSPAS TMAPIPSALAVWEPAGSSPQLSSAPADSS\PLP ALPKVLPPWTQKPLLGCLCQSPLPLLSPPDQV RCPPACSPAAASSFSFESQPCPSAPSKASPAPA ALIVQPHHPP*SQPQSQSVHPHGPGGPQPPL AASSLFWMFCQPPPPHPQFLWHRPLPVTGKA LASVLCFRPAPGSLRQTPLPPQFHIPRPGLSAP/ PPASGTSDSSDSRSPSASARVWPPAISPPPP AARHRPHPPEYFLSPCPFSCGFPRLLGRPRRPQ ALQTFRAWDLPPGSSPAPLCSGPELP*APPPLP PFPX*LGSGHPPSAQVPGLW*RCV*GHPIP RPVGHS*SGPPHSPL*APPQAWPLELPPSRQC LQPLHLRAAQPLDPCCSLSPPGPPLPVALPS WPGRP*SPSPASQPYHAGLPGPQSSPLPPGL PQLPSLRSGSQQPLLFFQCPGPGAV WGKGSPQ PLSPHPPPP/ARTQTFPVASRSLSPGTAPYSVCL TPSRSASSLPEVVLASSLPKIPQSSGSPLGFTSP MP*CFHRPSPPLLAQQHQPSPPSPHARSTLT	516	1006	<u> </u>	11.57			
FLVETGFLHYUGQASLELPTSGDTPAS\ASQSA GITGVSHHA*PRASGRCW 1897 A 4176 3029 1 AGPDGLAAPASCQGARGQTRVPGAFSWLAP GSHHASEGLAPGVPPAGGVSAQELTAPPQEG WGLGAPPAAPRPESDEKRAGSDAVRSFSRGA RDSLGQRRLGGTRGAGPAGKGAQRTMGPAS GFHSFPPRPHQEPSPRSSCWQHLLWHCPWPQ PSRLPRLTPAQLLQGPGVLAAPPGP*HVPGFL AQSPWPLPSGPRSP*DPLHQGALVPLPQGGSP HTAPHCLPSVLSPAIQQPLLPTAST/SSRSPPAS TMAPHCLPSVLSPAIQQPLLPTAST/SSRSPPAS TMAPHCLPSVLSPAIQQPLLPTAST/SSRSPPAS ALPKVLPPWTQKPLLGCLCQSPLPLLSPPDQI/ RCPPACSPAAASSFSFESQPCPSAPSKASPAPA ALIVGPHHPP*SQQPQSQSVHPHGPGGPQPPL AASSLFWMFCQPPPPHPQFLWHRPLPVTGKA LASVPLCFRPAPGSLRQTPLPPQFHHPRFGLSAP/ PPPASGTSDSSDSRSPSASAARVWPPA\SPPPP AARHRPHPEVTLSPCPFSCGFPRLLGRPRRPQ ALQTPRAWDLPPGSSPAPLCSGPELP*APPPLP PFPVA*LGSGHPPSAQVPGLW*RCV*GHPIP RPVGHS*SGPPHSPDL*APPQAWPLELPPSRQC LQPLHLRAAQPLDPCCSLSPPGPPLPVPALPS WPGRP*SPSPASSQPPYHAGLPGPQSSPLPPGL PPLPSLRSGSQQPLLFFQCPGPGAVWGKGSPQ PLSPHPPPP/ARTQTFPVASRSLSPGTAPYSVCL TPSRSASSLPEVVLASSLPKIPQSSGSVPLGFTSP MP*CFHRPSPPLP/LSSPFPAURPQAPQFPLHLP P*PPAPSPGCPLPPLAQQHQPSPPSPHARSTLT	346	1896	A	4174	1252	1190	FFQVFIFLFLIFFKTEFHSCCPGAVQWHDLDSL
GITGVSHHA*PRASGRRCW 547 1897 A 4176 3029 I AGPDGLAAPASCQGARGQTRVPGAFSWLAP GSHHASEGLAPGVPPAGGVSAQELTAPPQEG WGLGAPPAAPRESDEKRAGSDAVRSFSRGA RDSLGQRRLGGTRGAGPAGKGAQRTMGPAS GFHSFPPRPHQEPSPRSSCWQHLLWHCPWPQ PSRLPRLTPAQLLQGPGVLAAPPGP*HVPGFL AQSPWPLPSGPRSP*DPLHQGALVPLPQGGSP HTAPHCLPSVLSPAIQQPLLPTAST/SSRSPPAS TMAPIPSALAVWEPAGSSPQLSSAPADSSVPLP ALPKVLPPWTQKPLLGCLQSPLPLLSPPDQV RCPPACSPAAASSFFESQPCPSAPSKASPAPA ALVIVGPHHPP*SQQPQSSVHPHGPGGPQPPL AASSLFWMFCQPPPPHPGFLWHRPLPVTGKA LASPLCFRPAPGSLRQTPLPPQFHIPRPGLSAP/ PPPASGTSDSSDSRSPSASAARVWPPA\SPPPP AARHRPHPPEYFLSPCPFSCGFPRLLGRPRRPQ ALQTPRAV*LGSGHPPSAQVPGLW*RCV*GHPIP RPVGHS*SGPPHSPL*APPQAWPLELPPSRQC LQPLHLRAAQPLDPCCSLSPPGFPLPVALPS WPGRP*SPSPASSQPPYHAGLPGPQSSPLPPGL PQLPSLRSGSQQPLLFFQCPGPGAVWGKGSPQ PLSPHPPPP/ARTQTFPVASRSLSPGTAPYSVCL TPSRSASSLPEVVLASSLPKIPQSSGSVPLGFTSP MP*CFHRPSPPLP/LSSPFPAULRPQAPQFPLHLP P*PPAPSPGCPLPPLAQQHQPSPPLPHARSTLT							QPPPPRFKGFSCLSLPSSWDYRHAPAHPANFV
1897 A 4176 3029 1 AGPDGLAAPASCQGARGQTRVPGAFSWLAP GSHHASEGLAPGVPPAGGVSAQELTAPPQEG WGLGAPPAAPRPESDEKRAGSDAVRSFSRGA RDSLGQRRLGGTRGAGPAGKGAQRTMGPAS GFHSFPPRPHQEPSPRSSCWQHLLWHCPWPQ PSRLPRLTPAQLLQGPGVLAAPPGP*HVPGFL AQSPWPLPSGPRSP*DPLHQGALVPLPQGGSP HTAPHCLPSVLSPAIQQPLLPTAST/SSRSPPAS TMAPIPSALAVWEPAGSSPQLSSAPADSSVPLP ALPKVLPPWTQKPLLGCLQSPLPLLSPPDQI/ RCPPACSPAAASSFSFESQPCPSAPSKASPAPA AL\IVGPHHPP*SQQPQSQSVHPHGPGGPQPPL AASSLFWMFCQPPPPHPQFLWHRPLPVTGKA LAS\PLCFRPAPGSLRQTPLPPQFHIPRPGLSAP/ PPPASGTSDSSDSRSPSASAARVWPPAISPPPP AARHRPHPPEYFLSPCPFSCGFPRLLGRPRRPQ ALQTPRAWDLPPGSSPAPLCSGPELP*APPPLP PFPRVA*LGSGHPPSAQVPGLW*RCV*GHPIP RPVGHS*SGPPHSPPL*APPQAWPLELPPSRQC LQPLHLRAAQPLDPCCSLSPPGPPLPVPALPS WPGRP*SPSPASSQPPYHAGLPGPQSSPLPPGL PQLPSLRSGSQQPLLFFQCPGPGAVWGKGSPQ PLSPHPPPP/ARTQTTPVASRSLSPGTAPYSVCL TPSRSASSLPEVVLASSLPKIPQSSGS\PLEGTTSP MP*CFHRPSPPLP/LSSPFPALRPQAPQFPLHLP P*PPAPSPGCPLPPLAQQHQPSPPSPHARSTLT	 	 					
GSHHASEGLAPGVPPAGGVSAQELTAPPQEG WGLGAPPAAPRPESDEKRAGSDAVRSFSRGA RDSLGQRRLGGTRGAGPAGKGAQRTMGPAS GFHSFPPRPHQEPSPRSSCWQHLLWHCPWPQ PSRLPRLTPAQLLQGPGVLAAPPGP*HVPGFL AQSPWPLPSGPRSP*DPLHQGALVPLPQGGSP HTAPHCLPSVLSPAIQQPLLPTAST/SSRSPPAS TMAPIPSALAVWEPAGSSPQLSSAPADSSVPLP ALPKVLPPWTQKPLIGCLCQSPLPLLSPPDQI/ RCPPACSPAAASSFSFESQPCPSAPSKASPAPA AL\IVGPHHPP*SQQPQSQSVHPHGPGGPQPPL AASSLFWMFCQPPPPHPQFLWHRPLPVTGKA LASSPLCFRPAPGSLRQTPLPPOFHIPRPGLSAP/ PPPASGTSDSSDSRSPSASAARVWPPA\SPPPP AAARHRPHPPEVFLSPCPFSCGFPRLLGRPRRPQ ALQTPRAWDLPPGSSPAPLCSGPELP*APPPLP PFPRVA*LGSGHPPSAQVPGLW*RCV*GHPIP RPVGHS*SGPPHSPPL*APPQAWPLELPPSRQC LQPLHLRAAQPLDPCCSLSPPGPPLPVPALPS WPCRP*SPSPASSQPPYHAGLPGPQSSPLPPGL PQLPSLRSGSQQPLLFFQCPGPGAVWGKGSPQ PLSPHPPPP/ARTQTTPVASRSLSPGTAPYSVCL TPSRSASSLPEVVLASSLPKIPQSSGS\PLGPTSP MP*CFHRPSPPLP/LSSCPFPAULRPQAPQFPLHLP P*PPAPSPGCPLPPLAQQHQPSPPSPHARSTLT	547	1907		4126	2020		
WGLGAPPAAPRPESDEKRAGSDAVRSFSRGA RDSLGQRRLGGTRGAGPAGKGAQRTMGPAS GFHSFPPRPHQEPSPRSSCWQHLLWHCPWPQ PSRLPRLTPAQLLQGPGVLAAPPGP*HVPGFL AQSPWPLPSGPRSP*DPLHQGALVPLPQGGSP HTAPHCLPSVLSPAIQQPLLPTAST/SSRSPPAS TMAPIPSALAVWEPAGSSPQLSSAPADSSVPLP ALPKVLPPWTQKPLLGCLCQSPLPLLSPPDQI/ RCPPACSPAAASSFSFESQPCPSAPSKASPAPA ALIVGPHHPP*SQQPQSQSVHPHGPGGPQPPL AASSLFWMFCQPPPHPQFLWHRPLPVTGKA LAS\PLCFRPAPGSLRQTPLPPQFHIPRPGLSAP/ PPPASGTSDSSDSRSPSASAARVWPPA\SPPPP AARHRHPPEYFLSPCPFSCGFPRLLGRPRRPQ ALQTPRAWDLPPGSSPAPLCSGPELP*APPPLP PFPRVA*LGSGHPPSAQVPGLW*RCV*GHPIP RPVGHS*SGPPHSPPL*APPQAWPLELPPSRQC LQPLHLRAAQPLDPCCSLSPPGPPLPVPALPS WPGRP*SPSPASQPPYHAGLPGPQSSPLPFGL PQLPSLRSGSQQPLLFFQCPGPGAVWGKGSPQ PLSPHPPPPARTQTTFPVASRSLSPGTAPYSVCL TPSRSASSLPEVVLASSLPKIPQSSGSPLGPTSP MP*CFHRPSPPLP/LSSPFPAULRQAPQFPLHLP P*PPAPSPGCPLPPLAQQHQPSPPPPHALSTLT	347	1037	Α	41/0	3029	1	
RDSLGQRRLGGTRGAGPAGKGAQRTMGPAS GFHSFPRPHQEPSPRSSCWQHLLWHCPWPQ PSRLPRLTPAQLLQGPGVLAAPPGP*HVPGFL AQSPWPLPSGPRSP*DPLHQGALVPLPQGGSP HTAPHCLPSVLSPAIQQPLLPTAST/SSRSPPAS TMAPIPSALAVWEPAGSSPQLSSAPADSSVPLP ALPKVLPPWTQKPLLGCLCQSPLPLLSPPDQI/ RCPPACSPAAASSFSFESQPCPSAPSKASPAPA ALVIVGPHHPP*SQQPQSQSVHPHGPGGPQPPL AASSLFWMFCQPPPHPQFLWHRPLPVTGKA LASVPLCFRPAPGSLRQTPLPPQFHIPRPGLSAP/ PPPASGTSDSSDSRSPSASAARVWPPA\SPPPP AARHRPHPPEYFLSPCPFSCGFPRLLGRPRRPQ ALQTPRAWDLPPGSSPAPLCSGPELP*APPPLP PFPRVA*LGSGHPPSAQVPGLW*RCV*GHPIP RPVGHS*SGPPHSPPL*APPQAWPLELPPSRQC LQPLHLRAAQPLDPCCSLSPPGPPLPVPALPS WPGRP*SPSPASSQPPYHAGLPGPQSSPLPPGL PQLPSLRSGSQQPLLFFQCPGPGAVWGKGSPQ PLSPHPPPP/ARTQTFPVASRSLSPGTAPYSVCL TPSRSASSLPEVVLASSLPKIPQSSGS\PLGPTSP MP*CFHRPSPPLP/LSSPFPAURPQAPQFPLHLP P*PPAPSPGCPLPPLAQQHQPSPPSPHARSTLT							
GFHSFPPRHQEPSPRSSCWQHLLWHCPWPQ PSRLPRLTPAQLLQGPGVLAAPPGF*HVPGFL AQSPWPLPSGPRSP*DPLHQGALVPLPQGGSP HTAPHCLPSVLSPAIQQPLLPTAST/SSRSPPAS TMAPIPSALAVWEPAGSSPQLSSAPADSS/PLP ALPKVLPPWTQKPLLGCLCQSPLPLLSPPDQI/ RCPPACSPAAASSFSESQPCPSAPSKASPAPA ALIVGPHHPP*SQQPQSQSVHPHGPGGPQPPL AASSLFWMFCQPPPPHPQFLWHRPLPVTGKA LASIPLCFRPAPGSLRQTPLPPOFHIPRPGLSAP/ PPPASGTSDSSDSRSPSASAARVWPPASPPPP AARHRPHPPEYFLSPCPFSCGFPRLLGRPRRPQ ALQTPRAWDLPPGSSPAPLCSGPELP*APPPLP PFPRVA*LGSGHPPSAQVPGLW*RCV*GHPIP RPVGHS*SGPPHSPPL*APPQAWPLELPPSRQC LQPLHLRAAQPLDPCCSLSPPGPPLPVPALPS WPGRP*SPSPASSQPPYHAGLPGPQSSPLPPGL PQLPSLRSGSQQPLLFFQCPGPGAVWGKGSPQ PLSPHPPPP/ARTQTFPVASRSLSPGTAPYSVCL TPSRSASSLPEVVLASSLPKIPQSSGSVPLGPTSP MP*CFHRPSPPLP/LSSPFPAULRPQAPQFPLHLP P*PPAPSPGCPLPPLAQQHQPSPPSPHARSTLT							
PSRLPRLTPAQLLQGPGVLAAPPGP*HVPGFL AQSPWPLPSGPRSP*DPLHQGALVPLPQGGSP HTAPHCLPSVLSPAIQQPLLPTAST/SSRSPPAS TMAPIPSALAVWEPAGSSPQLSSAPADSS/PLP ALPKVLPPWTQKPLLGCLQSPLPLLSPPDQI/ RCPPACSPAAASSFSFESQPCPSAPSKASPAPA AL\IVGPHHPP*SQQPQSQSVHPHGPGGPQPPL AASSLFWMFCQPPPPHPQFLWHRPLPVTGKA LASIPLCFRPAPGSLRQTPLPPQFHIPRPGLSAP/ PPPASGTSDSSDSRSPSASAARVWPPA\SPPPP AARHRPHPPEYFLSPCPFSCGFPRLLGRPRRPQ ALQTPRAWDLPPGSSPAPLCSGPELP*APPPLP PFPRVA*LGSGHPPSAQVPGLW*RCV*GHPIP RPVGHS*SGPPHSPPL*APPQAWPLELPPSRQC LQPLHLRAAQPLDPCCSLSPPGPPLPVPALPS WPGRP*SPSPASSQPPLHFQCPGPGAVWGKGSPQ PLSPHPPPP/ARTQTFPVASRSLSPGTAPYSVCL TPSRSASSLPEVVLASSLPKIPQSSG\PLGPTSP MP*CFHRPSPPLPLSSPFPA\LRPQAPQFPLHLP P*PPAPSPGCPLPPLAQQHQPSPPSPHARSTLT		1 1		i i	İ	1	
AQSPWPLPSGPRSP*DPLHQGALVPLPQGGSP HTAPHCLPSVLSPAIQQPLLPTAST/SSRSPPAS TMAPIPSALAVWEPAGSSPQLSSAPADSS/PLP ALPKVLPPWTQKPLLGCLCQSPLPLLSPPDQI/ RCPPACSPAAASSFSESQPCPSAPSKASPAPA AL\IVGPHHPP*SQQPQSQSVHPHGPGGPQPPL AASSLFWMFCQPPPHPQFLWHRPLPVTGKA LASSPLCFRPAPGSLRQTPLPPQFHIPRPGLSAP/ PPPASGTSDSSDSRSPSASAARVWPPASPPPP AARHRPHPPEYFLSPCPFSCGFPRLLGRPRRPQ ALQTPRAWDLPPGSSPAPLCSGPELP*APPPLP PFPRVA*LGSGHPPSAQVPGLW*RCV*GHPIP RPVGHS*SGPPHSPPL*APPQAWPLELPPSRQC LQPLHLRAAQPLDPCCSLSPPGPPLPVPALPS WPGRP*SPSPASSQPPYHAGLPGPQSSPLPPGL PQLPSLRSGSQQPLLFFQCPGPGAVWGKGSPQ PLSPHPPPP/ARTQTFPVASRSLSPGTAPYSVCL TPSRSASSLPEVVLASSLPKIPQSSGS\PLGPTSP MP*CFHRPSPPLP/LSSPFPA\LRPQAPQFFLHLP P*PPAPSPGCPLPPLAQQHQPSPPSPHARSTLT	Ī	1					
HTAPHCLPSVLSPAIQQPLLPTAST/SSRSPPAS TMAPIPSALAVWEPAGSSPQLSSAPADSS\PLP ALPKVLPPWTQKPLLGCLCQSPLPLLSPPDQI/ RCPPACSPAAASSFSFESQPCPSAPSKASPAPA AL\IVGPHIPP*SQQPQSQSVHPHGPGGPQPPL AASSLFWMFCQPPPPHPQFLWHRPLPVTGKA LAS\PLCFRPAPGSLRQTPLPPQFHIPRPGLSAP/ PPPASGTSDSSDSRSPSASAARVWPPA\SPPPP AARHRPHPPEYFLSPCPFSCGFPRLLGRPRRPQ ALQTPRAWDLPPGSSPAPLCSGPELP*APPPLP PFPRVA*LGSGHPPSAQVPGLW*RCV*GHPIP RPVGHS*SGPPHSPPL*APPQAWPLELPPSRQC LQPLHLRAAQPLDPCCSLSPPGPPLPVPALPS WPGRP*SPSPASSQPPYHAGLPGPQSSPLPPGL PQLPSLRSGSQQPLLFFQCPGPGAVWGKGSPQ PLSPHPPPP/ARTQTFPVASRSLSPGTAPYSVCL TPSRSASSLPEVVLASSLPKIPQSSGS\PLGPTSP MP*CFHRPSPPLP/LSSPFPA\URPQAPQFFLHLP P*PPAPSPGCPLPPLAQQHQPSPPSPHARSTLT		1				١ ١	AUCDADI DECED CD# DDI NOCYT ADI BOCCED
TMAPIPSALAVWEPAGSSPQLSSAPADSS\PLP ALPKVLPPWTQKPLLGCLCQSPLPLLSPPDQI/ RCPPACSPAAASSFSFESQPCPSAPSKASPAPA AL\IVGPHHPP*SQQPQSQSVHPHGPGGPQPPL AASSLFWMFCQPPPPHPQFLWHRPLPVTGKA LAS\PLCFRAPGSLRQTPLPPQFHIPRPGLSAP/ PPPASGTSDSSDSRSPSASAARVWPPA\SPPPP AARHRPHPPEYFLSPCPFSCGFPRLLGRPRRPQ ALQTPRAWDLPPGSSPAPLCSGPELP*APPPLP PFPRVA*LGSGHPPSAQVPGLW*RCV*GHPIP RPVGHS*SGPPHSPPL*APPQAWPLELPPSRQC LQPLHLRAAQPLDPCCSLSPPGPPLPVPALPS WPGRP*SPSPASSQPPYHAGLPGPQSSPLPPGL PQLPSLRSGSQQPLLFFQCPGGAVWGKGSPQ PLSPHPPPP/ARTQTFPVASRSLSPGTAPYSVCL TPSRSASSLPEVVLASSLPKIPQSSGS\PLGPTSP MP*CFHRPSPPLP/LSSPFPA\LRPQAPQFPLHLP P*PPAPSPGCPLPPLAQQHQPSPPSPHARSTLT						:	
ALPKVLPPWTQKPLLGCLCQSPLPLLSPPDQI/ RCPPACSPAAASSFSFESQPCPSAPSKASPAPA ALUVGPHHPP*SQQPQSQSVHPHGPGGPQPPL AASSLFWMFCQPPPPHPQFLWHRPLPVTGKA LAS\PLCFRPAPGSLRQTPLPQFHIPRPGLSAP/ PPPASGTSDSSDSRSPSASAARVWPPA\SPPPP AARHRPHPPEYFLSPCPFSCGFPRLLGRPRRPQ ALQTPRAWDLPPGSSPAPLCSGPELP*APPPLP PFPRVA*LGSGHPPSAQVPGLW*RCV*GHPIP RPVGHS*SGPPHSPPL*APPQAWPLELPPSRQC LQPLHLRAAQPLDPCCSLSPPGPPLPVPALPS WPGRP*SPSPASSQPPLHGLGPGPQSSPLPPGL PQLPSLRSGSQQPLLFFQCPGPGAVWGKGSPQ PLSPHPPPP/ARTQTFPVASRSLSPGTAPYSVCL TPSRSASSLPEVVLASSLPKIPQSSGS\PLGPTSP MP*CFHRPSPPLP/LSSFFPA\LRPQAPQFPLHLP P*PPAPSPGCPLPPLAQQHQPSPPSPHARSTLT	1	1 1		1 1			
RCPPACSPAAASSFSFESQPCPSAPSKASPAPA AL\IVGPHHPP*SQQPQSQSVHPHGPGGPQPPL AASSLFWMFCQPPPPHPQFLWHRPLPVTGKA LAS\PLCFRPAPGSLRQTPLPPQFHIPRPGLSAP/ PPPASGTSDSSDSRSPSASAARVWPPA\SPPPP AARHRPHPPEYFLSPCPFSCGFPRLLGRPRRPQ ALQTPRAWDLPPGSSPAPLCSGPELP*APPPLP PFPRVA*LGSGHPPSAQVPGLW*RCV*GHPIP RPVGHS*SGPPHSPPL*APPQAWPLELPPSRQC LQPLHLRAAQPLDPCCSLSPPGPPLPVPALPS WPGRP*SPSPASSQPPYHAGLPGPQSSPLPPGL PQLPSLRSGSQQPLLFFQCPGPGAVWGKGSPQ PLSPHPPPP/ARTQTFPVASRSLSPGTAPYSVCL TPSRSASSLPEVVLASSLPKIPQSSGS\PLGPTSP MP*CFHRPSPPLP/LSSFFPA\LRPQAPQFPLHLP P*PPAPSPGCPLPPLAQQHQPSPPSPHARSTLT	1				İ		ALPKVLPPWTOKPLLGCLCOSPLPLLSPPDOI/
AL\IVGPHHPP*SQQPQSQSVHPHGPGGPQPPL AASSLFWMFCQPPPPHPQFLWHRPLPVTGKA LAS\PLCFRPAPGSLRQTPLPPQFHIPRPGLSAP/ PPPASGTSDSSDSRSPSASAARVWPPA\SPPPP AARHRPHPPEYFLSPCPFSCGFPRLLGRPRRPQ ALQTPRAWDLPPGSSPAPLCSGPELP*APPPLP PFPRVA*LGSGHPPSAQVPGLW*RCV*GHPIP RPVGHS*SGPPHSPPL*APPQAWPLELPPSRQC LQPLHLRAAQPLDPCCSLSPPGPPLPVPALPS WPGRP*SPSPASSQPPYHAGLPGPQSSPLPPGL PQLPSLRSGSQQPLLFFQCPGPGAVWGKGSPQ PLSPHPPPP/ARTQTFPVASRSLSPGTAPYSVCL TPSRSASSLPEVVLASSLPKIPQSSGS\PLGPTSP MP*CFHRPSPPLP/LSSPFPA\LRPQAPQFPLHLP P*PPAPSPGCPLPPLAQQHQPSPPSPHARSTLT	1						RCPPACSPAAASSESESOPCPSAPSKASPAPA
AASSLFWMFCQPPPPHPQFLWHRPLPVTGKA LAS\PLCFRPAPGSLRQTPLPPQFHIPRPGLSAP/ PPPASGTSDSSDSRSPSASAARVWPPA\SPPPP AARHRPHPPEVFLSPCPFSCGFPRLLGRPRRPQ ALQTPRAWDLPPGSSPAPLCSGPELP*APPPLP PFPRVA*LGSGHPPSAQVPGLW*RCV*GHPIP RPVGHS*SGPPHSPPL*APPQAWPLELPPSRQC LQPLHLRAAQPLDPCCSLSPPGPPLPVPALPS WPGRP*SPSPASSQPPYHAGLPGPQSSPLPPGL PQLPSLRSGSQQPLLFFQCPGPGAVWGKGSPQ PLSPHPPPP/ARTQTFPVASRSLSPGTAPYSVCL TPSRSASSLPEVVLASSLPKIPQSSGS\PLGPTSP MP*CFHRPSPPLP/LSSPFPA\LRPQAPQFFLHLP P*PPAPSPGCPLPPLAQQHQPSPPSPHARSTLT]				I		AL\IVGPHHPP*SOOPOSOSVHPHGPGGPOPPI
LAS\PLCFRPAPGSLRQTPLPPQFHIPRPGLSAP/ PPPASGTSDSSDSRSPSASAARVWPPA\SPPPP AARHRPHPPEYFLSPCPFSCGFPRLLGRPRRPQ ALQTPRAWDLPPGSSPAPLCSGPELP*APPPLP PFPRVA*LGSGHPPSAQVPGLW*RCV*GHPIP RPVGHS*SGPPHSPPL*APPQAWPLELPPSRQC LQPLHLRAAQPLDPCCSLSPPGPPLPVPALPS WPGRP*SPSPASSQPPYHAGLPGPQSSPLPPGL PQLPSLRSGSQQPLLFFQCPGFGAVWGKGSPQ PLSPHPPPP/ARTQTFPVASRSLSPGTAPYSVCL TPSRSASSLPEVVLASSLPKIPQSSGS\PLGPTSP MP*CFHRPSPPLP/LSSPFPA\LRPQAPQFFLHLP P*PPAPSPGCPLPPLAQQHQPSPPSPHARSTLT		}		}			AASSLFWMFCOPPPPHPOFLWHRPLPVTGKA
PPPASGTSDSSDSRSPSASAARVWPPA\SPPPP AARHRPHPPEYFLSPCPFSCGFPRLLGRPRRPQ ALQTPRAWDLPPGSSPAPLCSGPELP*APPPLP PFPRVA*LGSGHPPSAQVPGLW*RCV*GHPIP RPVGHS*SGPPHSPPL*APPQAWPLELPPSRQC LQPLHLRAAQPLDPCCSLSPPGPPLPVPALPS WPGRP*SPSPASSQPPYHAGLPGPQSSPLPPGL PQLPSLRSGSQQPLLFFQCPGFGAVWGKGSPQ PLSPHPPPP/ARTQTFPVASRSLSPGTAPYSVCL TPSRSASSLPEVVLASSLPKIPQSSGS\PLGPTSP MP*CFHRPSPPLP/LSSPFPA\LRPQAPQFFLHLP P*PPAPSPGCPLPPLAQQHQPSPPSPHARSTLT	Ι .				ł		LAS\PLCFRPAPGSLRQTPLPPOFHIPRPGLSAP/
AARHRPHPPEYFLSPCPFSCGFPRLLGRPRRPQ ALQTPRAWDLPPGSSPAPLCSGPELP*APPPLP PFPRVA*LGSGHPPSAQVPGLW*RCV*GHPIP RPVGHS*SGPPHSPPL*APPQAWPLELPPSRQC LQPLHLRAAQPLDPCCSLSPPGPPLPVPALPS WPGRP*SPSASSQPPYHAGLPGPQSSPLPPGL PQLPSLRSGSQQPLLFFQCPGPGAVWGKGSPQ PLSPHPPPP/ARTQTFPVASRSLSPGTAPYSVCL TPSRSASSLPEVVLASSLPKIPQSSGSVPLGPTSP MP*CFHRPSPPLP/LSSPFPA\LRPQAPQFFLHLP P*PPAPSPGCPLPPLAQQHQPSPPSPHARSTLT					[J	PPPASGTSDSSDSRSPSASAARVWPPA\SPPPP
ALQTPRAWDLPPGSSPAPLCSGPELP*APPPLP PFPRVA*LGSGHPPSAQVPGLW*RCV*GHPIP RPVGHS*SGPPHSPPL*APPQAWPLELPPSRQC LQPLHLRAAQPLDPCCSLSPPGPPLPVPALPS WPGRP*SPSPASSQPPYHAGLPGPGSSPLPPGL PQLPSLRSGSQQPLLFFQCPGPGAVWGKGSPQ PLSPHPPPP/ARTQTFPVASRSLSPGTAPYSVCL TPSRSASSLPEVVLASSLPKIPQSSGSVPLGPTSP MP*CFHRPSPPLP/LSSPFPAVLRPQAPQFFLHLP P*PPAPSPGCPLPPLAQQHQPSPPSPHARSTLT							AARHRPHPPEYFLSPCPFSCGFPRLLGRPRRPO
PFPRVA*LGSGHPPSAQVPGLW*RCV*GHPIP RPVGHS*SGPPHSPPL*APPQAWPLELPPSRQC LQPLHLRAAQPLDPCCSLSPPGPPLPVPALPS WPGRP*SPSPASSQPPYHAGLPGPQSSPLPPGL PQLPSLRSGSQQPLLFFQCPGPGAVWGKGSPQ PLSPHPPPP/ARTQTFPVASRSLSPGTAPYSVCL TPSRSASSLPEVVLASSLPKIPQSSGSVPLGPTSP MP*CFHRPSPPLP/LSSPFPAVLRPQAPQFPLHLP P*PPAPSPGCPLPPLAQQHQPSPPSPHARSTLT		j j				ļ	ALQTPRAWDLPPGSSPAPLCSGPELP*APPPLP
RPVGHS*SGPPHSPPL*APPQAWPLELPPSRQC LQPLHLRAAQPLDPCCSLSPPGPPLPVPALPS WPGRP*SPSPASSQPPYHAGLPGPQSSPLPPGL PQLPSLRSGSQQPLLFFQCPGPGAVWGKGSPQ PLSPHPPPP/ARTQTFPVASRSLSPGTAPYSVCL TPSRSASSLPEVVLASSLPKIPQSSGS\PLGPTSP MP*CFHRPSPPLP/LSSPFPA\LRPQAPQFPLHLP P*PPAPSPGCPLPPLAQQHQPSPPSPHARSTLT		1 1			[. [PFPRVA*LGSGHPPSAQVPGLW*RCV*GHPIP
LQPLHLRAAQPLDPCCSLSPPGPPLPVPALPS WPGRP*SPSPASSQPPYHAGLPGPQSSPLPPGL PQLPSLRSGSQQPLLFFQCPGPGAVWGKGSPQ PLSPHPPPP/ARTQTFPVASRSLSPGTAPYSVCL TPSRSASSLPEVVLASSLPKIPQSSGS\PLGPTSP MP*CFHRPSPPLP/LSSPFPA\LRPQAPQFPLHLP P*PPAPSPGCPLPPLAQQHQPSPPSPHARSTLT	1				Ì	l	RPVGHS*SGPPHSPPL*APPQAWPLELPPSROC
PQLPSLRSGSQQPLLFFQCPGPGAVWGKGSPQ PLSPHPPPP/ARTQTFPVASRSLSPGTAPYSVCL TPSRSASSLPEVVLASSLPKIPQSSGS\PLGPTSP MP*CFHRPSPPLP/LSSPFPA\LRPQAPQFPLHLP P*PPAPSPGCPLPPLAQQHQPSPPSPHARSTLT			1		İ		
PLSPHPPPP/ARTQTFPVASRSLSPGTAPYSVCL TPSRSASSLPEVVLASSLPKIPQSSGS\PLGPTSP MP*CFHRPSPPLP/LSSPFPA\LRPQAPQFPLHLP P*PPAPSPGCPLPPLAQQHQPSPPSPHARSTLT						ŀ	
TPSRSASSLPEVVLASSLPKIPQSSGS\PLGPTSP MP*CFHRPSPPLP/LSSPFPA\LRPQAPQFPLHLP P*PPAPSPGCPLPPLAQQHQPSPPSPHARSTLT		1 1			i	l	
TPSRSASSLPEVVLASSLPKIPQSSGS\PLGPTSP MP*CFHRPSPPLP/LSSPFPA\LRPQAPQFPLHLP P*PPAPSPGCPLPPLAQQHQPSPPSPHARSTLT						1	PLSPHPPPP/ARTQTFPVASRSLSPGTAPYSVCL
P*PPAPSPGCPLPPLAQQHQPSPPSPHARSTLT							TPSRSASSLPEVVLASSLPKIPQSSGS\PLGPTSP
						ļ	
PPLWPSLALLP*PLPPPPPVPSFSASLLCSLPAH	1		1	}			
	L						PPLWPSLALLP*PLPPPPPVPSFSASLLCSLPAH

SEQ ID NO: of nucl- cotide	SEQ ID NO: of peptide seq-	Met hod	SEQ ID NO: in USSN	Predicted beginning nucleotide location	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phonylalanine, G=Glycine, H=Histidinc, I=Isoleucine, K=Lysine, L=Leucine,
seq- uence	uence		09/496 914	correspondi ng to first amino acid residue of peptide sequence	corresponding to last amino acid residue of peptide sequence	M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						GTPASPGLGRSCLGKPQTLPWISFWPPSGRLA PGTWQPW/PVSPAPLSCLSAWDPWELPSPQPQ VCSTAELPTSCLLSSPGP\PAFQPPRFGCL*GPP GPPGLPPLQSSLSFPPPPPPVPQPPAPPALQWG LHLPGGRTK
548	1898	A	4180	2369	844	RIHREEDFQFILKGIARLLSNPLLQTYLPNSTK KIQFHQELLVLFWKLCDFNKVGQPRGALQGD GEQLPQ*PGGRDSVRLRGVGQSCPSLELSPLG PSPHP*KFLFFVLKSSDVLDILVPILFFLNDAR ADQSRVGLMHIGVFILLLLSGECNFGVRLNKP YSIRVPMDIPVFTGTHADLLIVVFHKIITSGHQ RLQPLFDCLLTIVVNVSPYLKSLSMVTANKLL HLLEAFSTTWFLFSAAQNHHLVFFLLEVFNNI IQYQFDGNSNLVYAIIRKRSIFHQLANLPTDPP TIHKALQRRRTPEPLSRTGSQGGAPPWRAPA PLPLQSQAPSRPVWWLLQALTS*PRSPRCQR MAPCGPWNLSPSRAWRMAARLRGSPARHGG SSGDRP/HSSASGQWSPTPEWVLSWKSKLPLQ TIMRLLQVLVPQVEKICIDKGLTDESEILRFLQ HGTLVGLLPVPHPILIRKYQANSGTAMWFRT YMWGVIYLRNVDPPVWYDTDVKLFEIQRV
549	1899	A	4191	858	321	LPWQRLGVLLSRGKMAVTGWLESLRTAQKT ALLQDGRRKVHYLFPDGKEMAEEYDEKTSE LLVRKWRVKSALGAMGQWQLEVGDPAPLG AGNLGPELIKESNANPIFMRKDTKMSFQWRIR NLPYPKDVYSVSVDQKERCIIVRTTNKKYYK KFSIPDLDRHQLPLDDALLSFA\TPTAP
550	1900	A	4192	1	1980	IRHTGSDIAGVCGWLLLSGPCGVGLDLDSRLL GASAMRRSEVLAEESIVCLQKALNHLREIWE LIGIPEDQRLQRTEVVKKHIKELLDMMIAEEE SLKERLIKSISVCQKELNTLCSELHVEPFQEEG ETTILQLEKDLRTQVELMRKQKKERKQE\LKL LQEQDQELC\EILCMPHYDIDSASVPSLEELNQ
551	1901	A	4194	3	1008	FRQHVTTLRETKASRREEF/VSSIKRQIILCME ELDHTPDTSFERDVVCEDEDAFCLSLENIATUL QKLLRQ\LEMQKSQNEAVCEG\LRTQI\RELW DRLQIPEEEREAVATIMSGSKAKVRK\ALQ\LE VDRLEELEKCKTMKKVIEAIRVELVQYWDQC FYSQEQRQAFAPFCAEDYTESLLQLHDAEIVR LKNYYEVHKELFEGVQKWEETWRLFLEFER KASDPNRFTNRGGNLLKEEKQRAKLQKMLP KLEEELKARIELWEQEHSKAFMVNGQKFME YVAEQWEMHRLEKERAKQERQLKNKKQTET EMLYGSAPRTPSKRRGLAPNTPGKARKLNTT TMSNATANSSIRPIFGGTVYHSPVSRLPPSGSK PVAASTCSGKKTPRTGRHGANKENLELNGSI LSGGYPGSAPLQRNFSINSVASTYSEFADPSLS DSSTVGLQRELSKASKSDATSGILNSTNIQS AWHEGLVSSPAIGAYLSASYGDSLVVLVATV
	1201	Λ.	+174	J	1006	WALDICFILVAVPESLPEKMRPVSWGAQISW KQADPFASLKKVGKDSTVLL\ICITVCLSYLPE AG\QYSSFF\LYLR\QVIGFG\TVKIAAFIAMVGI LSIVAQTAFLSILMRSLGNKNTVLLGLGFQML QLAWYGFGSQAWMMWAAGTVAAMSSITFP AISALVSRNAESDQQGVAQGIITGIRGLCNGL GPALYGFIFYMFHVELTELGPKLNSNNVPLQ GAVIPGPPFLFGACIVLMSFLVALFIPEYSKAS GVQKHSNSSSGSLTNTPERGSDEDIEPLLQDS SIWELSSFEEPGNQCTEL

1	SEQ ID	SEQ ID	Met	SEQ	Dradicted	Deadigted and	Amino polidocuro (A-11: 1: C.O. 11
	NO: of	NO: of	hod	ID NO:	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
ı	nucl-	peptide	1100	in NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, I=Histidine,
1	cotide	seq-		USSN	location	corresponding	
ı	seq-	uence		09/496	correspondi	to last amino	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
- 1	uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
١		J		1 717	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
ſ					residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
					peptide	Sequence	/=possible nucleotide deletion, \=possible
- 1					sequence		nucleotide insertion
ł	552	1902	A	4197	2	14302	ARPPPAPGSRQQKQKAAPGAAAAAELRGAR
	-		-	j	-	1.502	EPAPARRGTMADGGEGEDEIOFL'RTDDEVV
ı						·	LQCTATIHKEQQKLCLAAEGFGNRLCFLESTS
- [i i			NSKNVPPDLSICTFVLEQSLSVRALQEMLANT
1					i		VEKSEGQVDVEKWKFMMKTAQGGGHRTLL
-							YGHAILLRHSYSGMYLCCLSTSRSSTDKLAFD
ſ							VGLQEDTTGEACWWTIHPASKQRSEGEKVR
- 1				}			VGDDLILVSVSSERYLHLSYGNGSLHVDAAF
-							QQTLWSVAPISSGSEAAQGYLIGGDVLRLLH
ł							GHMDECLTVPSGEHGEEQRRTVHYEGGAVS
- 1]		VHARSLWRLETLRVAWSGSHIRWGQPFRLR
-							HVTTGKYLSLMEDKNLLLMDKEKADVKSTA
1			:				FTFRSSKEKLDVGVRKEVDGMGTSEIKYGDS
J							VCYIQHVDTGLWLTYQSVDVKSVRMGSIQR
-					İ	, i	KAIMHHEGHMDDGISLSRSQHEESRTARVIRS
1		Ì					TVFLFNRFIRGLDALSKKAKASTVDLPIESVSL
١							SLQDLIGYFHPPDEHLEHEDKQNRLRALKNR
ı							QNLFQEEGMINLVLECIDRLHVYSSAAHFAD
1				}			VAGREAGESWKSILNSLYELLAALIRGNRKN
							CAQFSGSLDWLISRLERLEASSGILEVLHCVL
							VESPEALNIKEGHIKSIISLLDKHGRNHKVLD
1					ł		VLCSLCVCHGVAVRSNQHLICDNLLPGRDLL LQTRLVNHVSSMRPNIFLGVSEGSAQYKKWY
							YELMVDHTEPFVTAEATHLRVGWASTEGYSP
							YPGGEEWGGNGVGDDLFSYGFDGLHLWSG
ł					1		CIARTVSSPNQHLLRTDDVISCCLDLSAPSISF
							RINGQPVQGMFENFNIDGLFFPVVSFSAGIKV
ļ							RFLLGGRHGEFKFLPPPGYAPCYEAVLPKEKL
1		Ì		1			KVEHSREYKQERTYTRDLLGPTVSLTQAAFT
J		J		}			PIPVDTSQIVLPPHLERIREKLAENIHELWVMN
-	i						KIELGWQYGPVRDDNKRQHPCLVEFSKLPEQ
					_		ERNYNLQMSLETLKTLLALGCHVGISDEHAE
1]				J	1	DKVKKMKLPKNYQLTSGYKPAPMDLSFIKLT
					1		PSQEAMVDKLAENAHNVWARDRIRQGWTY
]	ļ		ŀ		1	GIQQDVKNRRNPRLVPYTPLDDRTKKSNKDS
	J	J	j]	j		LREAVRTLLGYGYNLEAPDQDHAARAEVCS
1	1		ļ	ĺ		[GTGERFRIFRAEKTYAVKAGRWYFEFETVTA
	Ì				į		GDMRVGWSRPGCQPDQELGSDERAFAFDGF
-	J	J	1	J			KAQRWHQGNEHYGRSWQAGDVVGCMVDM NEHTMMFTLNGEILLDDSGSELAFKDFDVGD
ĺ		1	i		· ·	į	GFIPVCSLGVAQVGRMNFGKDVSTLKYFTIC
1	ł	}	ł	}	ł		GLQEGYEPFAVNTNRDITMWLSKRLPQFLQV
	ļ		ĺ	ļ		ŀ	PSNHEHIEVTRIDGTIDSSPCLKVTQKSFGSQN
	į					1	SNTDIMFYRLSMPIECAEVFSKTVAGGLPGAG
	ł	1	ł	1	1	i	LFGPKNDLEDYDADSDFEVLMKTAHGHLVP
		1					DRVDKDKEATKPEFNNHKDYAQEKPSRLKQ
		ļ	į			ĺ	RFLLRRTKPDYSTSHSARLTEDVLADDRDDY
1		1	- 1	ł			DFLMQTSTYYYSVRIFPGQEPANVWVGWITS
				- 1			DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE
							SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC
1		- 1	1	1	1	1	VVDAASGLLTFIANGKELSTYYQVEPSTKLFP
	ļ	ļ			İ		AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS
	- 1	ĺ	- 1	1	İ		EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL
1		1	l				KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN
		-	F				RSVDILELTEQEELLKFHYHTLRLYSAVCALG
	i	. !	1	ľ	- 1		NHRVAHALCSHVDEPQLLYAIENKYMPGLLR
		İ	·		!		AGYYDLLIDIHLSSYATARLMMNNEYIVPMT
	i i			1	j		EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF
	i			1	J.		
1	1		ł	İ			SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE AVKEGSLHARDPVGGTTEFLFVPLIKLFYTLLI

SEQ ID NO: of NO: of NO: of nucl- eotide seq- uence uence	utamic Acid, ycine, H=Histidine, L=Leucine, aragine, P=Proline, ine, S=Serine,
nucl- cotide seq- uence uence	ycine, H=Histidine, L=Leucine, aragine, P=Proline, ine, S=Serine,
eotide seq- seq- uence USSN location corresponding to last amino M=Methionine, N=Aspa uence 914 ng to first acid residue Q=Glutamine, R=Argini	L=Leucine, aragine, P=Proline, ine, S=Serine,
seq- uence 09/496 correspondi to last amino M=Methionine, N=Aspa uence 914 ng to first acid residue Q=Glutamine, R=Argini	aragine, P=Proline, ine, S=Serine,
uence 914 ng to first acid residue Q=Glutamine, R=Argini	ine, S=Serine,
ing to first actual residue Q=Olutainine, K=Argini	ine, o-ocinie,
,	W-Tayatonbon
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	
	ieiion, \=possible
sequence nucleotide insertion	
	IEPSVFKEAATPEEESDT
	AGEEEAKGGKRPKEGLL
	LLQYLCDCQVRHRIEAI
	QRFRYNEVMQALNMSA
	EQINMLLNFKDDKSECP
CPEEIRDQLLDFHEDI	LMTHCGIELDEDGSLDG
	VTYLKKKQAEKPVES
DSKKSSTLQQLISETM	AVRWAQESVIEDPELVR
	GLVRALPKTYTINGVSV
	LSVRMGKEEEKLMIRG
LGDIMNNKVFYQHPN	NLMRALGMHETVMEV
	PKMVANCCRFLCYFCR
ISRQNQKAMFDHLSY	LLENSSVGLASPAMRG
STPLDVAAASVMDN	NELALALREPDLEKVVR
	SKGYPDIGWNPVEGER
YLDFLRFAVFCNGES'	VEENANVVVRLLIRRPE
CFGPALRGEGGNGLL	.AAMEEAIKIAEDPSRD
GPSPNSGSSKTLDTER	EEEDDTIHMGNAIMTFY
SALIDLLGRCAPEMH	LIHAGKGEAIRIRSILRS
LIPLGDLVGVISIAFQI	MPTIAKDGNVVEPDMS
AGFCPDHKAAMVLFI	LDRVYGIEVQDFLLHLL
EVGFLPDLRAAASLD	TAALSATDMALALNRY
LCTAVLPLLTRCAPLI	FAGTEHHASLIDSLLHT
VYRLSKGCSLTKAQR	DSIEVCLLSICGQLRPS
MMQHLLRRLVFDVP	LLNEHAKMPLKLLTNH
	GNFGAASEEELHLSRK
	QELFKLALPCLSAVAG
	MEKOSSMDSEGNFNPO
	NKYAEHSHDKWSMDK
	KVQPLMKPYKLLSEKE
, , , , , , ,	LARTMRTERTREGDSM
	SVDAAHGYSPRAIDMS
	MAENYHNIWAKKKKM
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	YDTLTAKEKAKDREKA
	RGFKDLELDTPSIEKRFA
	QYILEFDGGSRGKGEHF
	IDQYFKNHRLYFLSAA
	EMVTSLFCKLGVLVRH
· · · · · · · · · · · · · · · · · · ·	HILGOTLDARTVMKTG
LESVKSALRAFLDNA	AEDLEKTMENLKQGQF
	(TTVALLPMLSSLFEHI
	VSCYRILTSLYALGTSK
	AFAGAFPVAFLETHLD
	AALSLPTNVEDVCPNIP
SLEKLMERIVELARSG	IRYTQMPHVMEVILPM
	NNPERAEMCCTALNSE
	LGIDEGAWMKRLAVF
	FLPLMEKLKKKAATVV
	BEAELLILDEFTTLARDL
1 1 1 1 1 1 1	AKWLKEPNPEAEELFR
MVAEVFIYWSKSHNF	
MODI INDTRODUACIONA	AVSDQERKKMKRKGD
DVOMOTOL TILA AT TOD	LLPIGLNICAPGDQELIA
K I SWIQI SLIV AALKII	DIDENIU OCE POP
LAKNRFSLKDTEDEVI	
AIRWQMALYKDLPNR	
DIANVLEHLEQKSKR	VGRRHYCLVEHPQRSK
KAV WHKLLSKQRKRA	AVVACFRMAPLYNLPR
HRAVNLFLQGYEKSW	/IETEEHYFEDKLIEDLA
KPGAEPPEEDEGTKR\	DIMARCHEREDES
EKCKLEEDFLYMAYA	DIMAKSCHDEEDDDG

NO: of not- ordinary period coation nucleotide coation of the coat		SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucle cotide scide cotide cotide contesponding of corresponding of corresponding unnoc unnoce when the contest of the cotide service of peptide residue of peptide sequence pept				1		1		D=Aspartic Acid, E=Glutamic Acid.
seq. uence 09/496 gridde amino acid residue of peptide amino acid residue of peptide sequence peptide	1	nucl-	peptide	ĺ	in	nucleotide	location	
uence 914 ng to first anino acid of applied residue of peptide residue of peptide sequence very	- [-	eotide	seq-	j	USSN	location	corresponding	
amino acid residue of peptide sequence T=Threonine, V=Valine, W=Tryptophan, residue of peptide sequence Sequence	1	seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
residue of peptide sequence Y=Tyrosine, X=Unknown, =*Sup codon, y-possible nucleotide deletion, 1-possible nucleotide deleti		uenœ			914			
Poptide sequence	ı			ļ	j			
sequence	1			1			sequence	
EEEWKSFEEKEMEKQKLLYQQARLHBRGAA EMM_OTINASKGETGPMVAATILGIAILM NSTVQQKMLDYLKEKKDVGFFQSLAGLMQ CSVLDINAFFERQNKAEGLGMYTEGGSGEK LQDDEFTCDLFRFLQLLCEGINSDFQNYLK QTGINNTTYNIIISTVDYLLRVQESISDFYWYY SGKDVIDEGGGGRKESKAIQVAKQVFNILTE QQPCTGNQOSLAHSSLWDAVVGFLHVFAH QMKLSQOSSQIELLKELMDLQKDMYVMLL MLEGRVVNGTIGKQMYDMLVESSNNVEMI KFPDMPLKLKDLTSSDTFKEVPDPGKGVISK RDFHKAMESIKHYTQSETEFLLSCAETDEN TLDYEFVKRPHEPAKDIGNYAVLUTINLSS MPNDTRLQTILLESVLNYFOPFLGRIEMS SAKRIERVYFEISESSTTOWERVQKESKRQ GIESDLNERSANKESSKEPPEGOPPMAF SILTVESALFALRYNDTLINRMLSLKSLKKO KVKKKMTVKDMYATFSSYWEISENTLLHFF ASVFRGFFRIICSLLLGGSLVEGAKKIKVAEL ANMPDPTQDEVRGDGEGGRKPLEAALPSE LTDLKETESDLLASSISTOMENPAPEVQEKFOCKAA SEKREKEPKSDESSERFANKENIAA QKLLNYFARRYNYMRMLALFYAFANFILL HNPNAGLSDLMSNPVPMPEVQEKFOCKAA EERKEEKEETKSEPEKERFAGGGOFKL HNPNAGLSDLMSNPVPMPEVQEKFOCKAA GKKKOKLRQLHTHRYGEPEVPESAFWKKIIAA QKLLNYFARRYNYMRMLALFYAFANFILL FYKVSTSSVVGGKELFTRSSSENAKVTSLDS SHRIIAVHYVLESSGVMEPTVRIJLPHTVIS FCIIGYYCLKVPLVFKREKEVARKLEFFOOL TEQPSEDDIGGQWRILVNTGSPFNYWDK VKRKVMDKYGGFYGRRRISHLGHMOKAAL FSDARREKKPKKDSSLSAVNINSVKYQMW VKRKVMDKYGGFYGRRRISHLGHMOKAAL SPDARREKKPKKDSSLSAVNINSVKYQMW KLGVVFTDNSFLYLAWYMT TEQESEDDIGGQWRILVNTGSPFNYWDK VKRRVMDKYGGFYGRRRISHLGHMOKAAL SPDARREKKPKSSLSAVNINSVKYQMW KLGVVFTDNSFLYLLAWYMT SPDARREKKPKSSLSAVNINSVKYQMW KLGVVFTDNSFLYLLAWYMT SPDARREKKPKSSLSAVNINSVKYQMW KLGVVFTDNSFLYLLAWYMT SPDARREKKPKSSLSAVNINSVKYQMW KLGVVFTDNSFLYLLAWYMT SPDARREKKPKSSLSAVNINSVKYQMW KLGVVFTDNSFLYLLAWYMT SPDARREKKPKPKSSLSSLSAVNINSVKYQMW KLGVVFTDNSFLYLLAWYMT SPDARREKKPKPKSSLSSLSAVNINSVKYQMW KLGVVFTDNSFLYLLAWYMT SPDARREKKPKPKSSLSSLSAVNINSVKYQMW KLGVVFTDNSFLYLCGLEVITELANGAR AVGLYLRIDKGRQGTRERFSERGGGRETQOVA LPLSHGRGGGGCCAAERVGAARGSAA AYGLYLRIDKGRCCCAREFYGGAGSTETTQOVA LPLSHGRGGGGCCAAERVGAARGSAA AYGLYLRIDKGRCCCAREFYGGGGSETTQOVA LPLSHGRGGGGCCAACAGRSCACACACACACACACACACACACACACACACACACACA				ł		,	}	
EMVLQTISASKGETGPWATIKLGIAILM NSTVQQKM.DYLKEKKDVGFFQSLGIMQ CSVLDINAFERQNKAEGLGMVTEEGSGLGIMQ CSVLDINAFERQNKAEGLGMVTEEGSGE LQDDEFTCDLFERLQLLCEGHRISPGNNYLR QGRCTGNQSLAHSRI,WDAVVGEILHFY SGKDVIDEQGGNFSKAIQVAKQVPNTLTF QGPCTGNOQSLAHSRI,WDAVVGEILHFY SGKDVIDEQGGNFSKAIQVAKQVPNTLTF QGPCTGNOQSLAHSRI,WDAVVGEILHFY QGPCTGNOQSLAHSRI,WDAVVGEILHFY QGPCTGNOQSLAHSRI,WDAVVGEILHFY QGPCTGNOQSLAHSRI,WDAVVGEILHFY QGPCTGNOQSLAHSRI,WDAVVGEILHFY QGPCTGNOQSLAHSRI,WDAVVGEILHFY QGPCTGNOQSLAHSRI,WDAVVGEILHFY QGPCTGNOQSLAHSRI,WDAVVGEILHFY QGPCTGNOQSLAHSRI,WDAVVGEITHS RPFHKAMESHKHYTQSSTEEFLISCAETDEN RAFFIELSCAETDEN RFFHKAMESHKHYTQSSTEEFLISCAETDEN RFFHKAMESHKHYTQSSTEEFLISCAETDEN RFFHKAMESHKHYTQSSTEEFLISCAETDEN RAFFIELSCAETDEN SAKRIERVYFEISESSRTQWEKPOPUGEKGE GFVVVNEGGEKEKMELFVNRCEPTIFFMQLA QISESDLNERSANKEESEKERPEQGFRAKI SAKRIERVYFEISESSSRTQWEKPOPUGEKGEGKOK KKVKKMTVKDMVTAFFSSYWSIFMTLHFH ASVFRGFFRICISCALLGGSLVGGKKIKVAEL ANMPPTQDEVRGDGEGGERKPLEAALPSE LTDLKEITEESDLASFIGLUSGGKEKKAAEL EEKEEKEETKSETKSEPEKAGEGIGKEKAAEL EEKEEKEETKSETKSEPEKAGEGIGKEKAAEL EKEKEEKEETKSETKSEPEKAGEGIGKEKAAEL EEKEEKEETKSETKSEPEKAGEGIGKEKEAAEL EEKEEKEETKSETKSEPEKAGEGIGKEKAAEL ANMPPTYMMTMALATVAFADIFIL FFKVSTSSVVGGKELPTRSSSENAAEVTSLDS SIRBILAHTVLEESSGFMETFYNWTM VKUKLWAMDKYGEFVERBARELLHINTUS FFIGIGYVCLKVPLVIFKREKEVAAKLEFDON VKUKLWAMDKYGEFVERBARELLHINTUS FFIGIGYVCLKVPLVIFKREKEVAAKLEFDON VKUKLWAMDKYGEFVERBARELHHTVS FFIGIGYVCLKVPLVIFKREKEVAAKLEFTON VKUKLWAMDKYGEFVERBARELHHTVS FFIGIGYVCLKVPLVIFKREKEVAAKLEFTON VKUKLWAMDKYGEFVERBARELHHTVS FFIGIGYTELMOON VKUKLWAMDKYGEFVERBAREHLHHTVS FFIGIGYTELMOON VKUKLWAMDKYGEFVERBAREHLHHTVS FFIGIGYTELMOON VKUKLWAMDKYGEFVERBAREHLHHTVS FFIGIGYTELMOON VKUKLWAMDKYGEFVERBAREHLHHTVS FFIGIGYTELMOON VKUKLWAMDKYGEFVERBAREHLHHTVS FFIGIGYTELMOON VKUKLWAMDKYGEFVERBAREHLHHTVS FFIGIGYTELMOON VKUKLWAMDKYGEFVERBAREHLHTVS FFIGIGYTELMOON VKUKLWAMDKYGEFVERBAREHLHTVS FFIGIGYTELMOON VKUKLWAMDKYGEFVERBAREHLHTVS FFIGIGYTEMOON VKUKLWAMDKYGEFVERBAREHLHTVS FFIGIGYTEMOON VKUKLWAMDKYGEFVERBAREH FFIGITSEELQOHLOON VKUKLWAMDKYGEFVERSHEDSILLE FFIGITSEELQOHLOON VKUKLWAMDKYGEFVERSHEDSILLE	<u> </u> _			<u> </u>	L	sequence		
NSTVÖÖKMLDYLKEKKDVGFFQSLAGUC CSVLDINAFERQNKAGCIGMVTEGSGEK LQDDEFTCDLFRIQILCEGHNSDFONYIR QITGNTTVNIISTVÖYLLRVÖESBFV MY- SGKDVIDEQGQRNFSKAIQVAKQVFNILTE QQFCTGQQSLAHSRI, WDAVVGFILVFNILTE QQFCTGQQSLAHSRI, WDAVVGFILVFNILTE QQFCTGQQSLAHSRI, WDAVVGFILVFNILTE QQFCTGQQSLAHSRI, WDAVVGFILVFNILTE QQFCTGQQSLAHSRI, WDAVVGFILVFNILTE QQFCTGQQSLAHSRI, WDAVVGFILVFNILTE QQFCTGQQSLAHSRI, WDAVVGFILVFNILTE MEGNVVNOTIGKQMVDMLVESSNNOWED REPMFAKKEN DE STENDEN								
CSVLDINAFERQNKAEGLGMVTEEGGER LQDDEFTCLIFERLQLCEGHINSPGNYLK QGTONNTTVNIIISTVDYLLRVOESISDFY SGKDVIDEGGGRYPSKAIQVAKQVPNTLYF GGPCTONQOSLAHSRLWDAVVGFLHVFAH QMKLSQDSSQIELLKELMDLQKDMVVMLI MEGGNVNOTTIGKQMVDMLVSSNNNEM KFFDMFLKIKDLTSSDTFKEVDPDGKGVJ RDFHAMESHKHYQSETEFLSCAETDEN TLDYEEFVKRHEPAKDIGFNVAVLLTING SAKRIERVYFEISESSTOWERPOYKESKRO FDVVNEGGEKSKMELFVNTCFIFEMQLA QISESDLNERSANKEESEKERPEGGFRMAF SILTVRSALFALRVIILTIMRNI,SKISLKO KKVKKMTVKDMVTAFFSSYWSIPMTILLIFF ASVPRGFFRIICSLLGGSLVEGAKKIKVAL HNPNAGLSDLMSNPVPMPEVOEKFQEKKL HNPNAGLSDLMSNPVPMPEVOEKFQEKKL EEKEEKETKSEPEKAGGDGEKEEKAKL KKGKLRQLHFIRVGEEPVPSSAFWKKIIA QQKLINYFARNFYMMMALFVAFAINFIL FYKVSTSSVVGEKLPTRSSSEAVSTLDS SHRIIAVHVLEESSGYMFETVRILPHITTINF FCIGGYYCLKVPLIFKREKEVAKVISLDS SHRIIAVHVLEESSGYMFETVRILPHITTINF FCIGGYYCLKVPLIFKREKEVAKLEFOGL TGPSEDDIKGOMDRI VINTOSPPNNY WDK VKUKNAMEVGERFERSHERHEV FROM VKUKNAMEVSGERFERSHERHEV FROM VKUKNAMEVSGERFERSHERHEV FROM VKUKNAMEVSGERFERSHERHEV FROM FROM FROM FROM FROM FROM FROM FROM	1				1	ļ	i	
LQDDEFTCDLFRICALCEGHNSDFONTUR. OTGNITTANIESTVDYLLINVESSIDFY WY SGKDVIDEQGORNFSKAJQVAKQVFINTLET QCPCTGNQQSLAHSRLWDAVVGFLHVFAH QMKLSQDSSQIELLKELMDLQKDMVVMLL MLEGNVVNOTTIGKQMVDMLVESSINNUGH KFFDMFLKLKDLISSDITKEVPDPGGKGVIS RDFHKAMESHKHYTQSETIEFLLSCAETDEN TILDYBEFVKRFHEPAKDIGFNVLSTNIE MPHOTRILOTFLELASSVLNYPQPFLGRIEM SAKRIERVYFEISESSRTOWERVESKRQ FDVVNEGGEKEKMELFVNCEDTIFEMQLA QISESDLNERSANKEESEKERPEEQGFRAML SAKRIERVYFEISESSRTOWERVESKRQ FDVVNEGGEKEKMELFVNCEDTIFEMQLA QISESDLNERSANKEESEKERPEEQGFRAML SAKRIERVYFEISESSRTOWERVESKRQ KKVKKNTVKDMVTAFFSSYWSIFMILLHY ASVYRGFFRICSLLGGGIVAGAKKIKVAEL ANMPDPTODEVRGDGEGERKPLEAALPI ANMPDPTODEVRGDGEGEGERKPLEAALPI ANMPDPTODEVRGDGEGEGERKPLEAALPI ANMPDPTODEVRGDGEGEGERKPLEAALPI ANMPDPTODEVRGDGEGERKPLEAALPI CEGRICAL COLLING ANMPDPTODEVRGGERGGARCA ANGLYLRIBKGRLQCLNESSEGGGARTGOQA ANGLYLRIBKGRLQCLNESSEGGGARTGOQA ANGLYLRIBKGRLQCLNESSEGGROVERPY ERADDRSKFVESDADEELLFNIPTGHHYCHT BERDGGGGGCAAARRVGAARSAA ANGLYLRIBKGRLQCLNESSEGGROVERPY ERADDRSKFVESDADEELLFNIPTGHTHYCH BERDGGGGGCCAAARRVGAARSAA ANGLYLRIBKGRLQCLNESSEGGROVERPY ERADDRSKFVESDADEELLFNIPTGHTHYCH BERDGGGGGCCAAARRVGAARSAA ANGLYLRIBKGRLQCLNESSEGGROVERPY ERADDRSKFVESDADEELLFNIPTGHTHYCH BERDGGGGGCCAAARRVGABSAACHA BERDGGGGGCCAAARRVGABSAACHA BERDGGGGGCCAAARRVGABRACHABDYGA BERDGGGGGCCAAARRVGABSAACHA BERDGGGGGGCGCAAARRVGABSAACHABDYGA BERDGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	-		i i		1			
QTGNNTTVNIIISTVDYLLRVQESISDFYNTITE QGRCTGNQOSLAHSRI, WDAWVGFLHYMY GOKDVIDEQGGNFSKAIQVAKQVFNTLTE QGRCTGNQOSLAHSRI, WDAWVGFLHYMYMIL MEGNVVNGTIGKQMAVMALL MLEGNVVNGTIGKQMAVDMLYSSSNNVEMI KFFDMFLKILKDLASVLNTYOPSIGETLISCAETDEN TLDYBEFVKRPHEPAKDIGFNVAVLLTING BOPPHRAMESHKHYTQSSTEEFLISCAETDEN TLDYBEFVKRPHEPAKDIGFNVAVLLTING SAKRIBEVYFEISESSRTQWEKPOVKESKRQ FDVVNEGGEKEKKMELFVNTCFIFEMQLA QISESDLNERSANKEESEKERPEGGFRAKE SILTVRSALFALRYNIILTIMRNTJENFYDFLGRIEM SAKRIBEVYFEISESSRTQWEKPEGVGKESKRQ GFVVVNEGGEKEKKMELFVNTCGAKKKVAEL ANMPDPTODEVRGDGEGERKPLEAALPSE LTDLKELTEESDLLSDIFGLDLKREGGGYVE ANMPDPTODEVRGDGEGERKPLEAALPSE LTDLKELTEESDLLSDIFGLLKREGGYVEK EEKEEKETKSEPFKAGGEDGEKEEKAKLIA QOKLLNYFARNFYNMRMLALFVAFAINFILL FYKVSTSSVVGGKEIPTRSSSFKAKVISLDS SHRIIAVHYLEESSGYMEPTVRILPILITITY FCIGGYYCLKVPLIVFKREKEVARKLEFDGL TEQPSEDDIKGQWDRI.VINTQSPPNNYWDK VKRKVAMDKVGGFFFRGRBREIHGNBAAL FSDAREKKKPKKDSSLSAVI.NSIDVKYQMW KLGVVFTDNSFLYLAWYMT 553 1903 A 4199 31 767 LPELNGRGAGLRRAEPSERGGGAERTQOVA LPLSHGHSHGGGCCRCAAERVGAARGSAA AYGLYLRINKGRIPTGAWAYMT LPELNGRGAGLRRAEPSERGGGAERTQOVA LPLSHGHSHGGGCCRCAAERVGAARGSAA AYGLYLRINKGRIQCLNESSEGROVFKPP ERADDRSKFVESDADEELLFNPFTGHVEL GIIMGEDDDSHSPEMRLYKNIPMSPDTIVEL HISKNFGADTTKYPYIGLRGEWITPLA WDFEDSTIQSFRWHPLRAKANKYEDSYPGS THE FPOTTSLNDLTGELEYATKISRFSNYYLL SICKACKNLAYGEEKKKCCNPVVKTYLLPI RSSQGKRKTGVQNNTUDPIFGETLKYQME QLVTRQLUSVWHLGTLARKPLGEVIPIPLA WDFEDSTIQSFRWHPLRAKANKYEDSYPGS TRAQLRQSSLELTVWDQALFGMNDRLLGGT RLGSKGDTAVQGALGAKNLPVRPGTLNSFVKGS LPTQALRQSSLELTVWDQALFGMNDRLLGGT RLGSKGDTAVQGALGAKNLPVRPGTLNSFVKGS TRAQLRQSSLELTVWDQALFGMNDRLLGGT RLGSKGDTAVQGALGAKNLPVRPSSLEDOLLE S555 1905 A 4211 331 2419 KENKKARNLRMNGSRSRSDGSEETLPQDH NHHENERR WQQELHREEAYYQFINEINDE DYRAMROHNLLGTGETISEEQQKLDOVL SPR								CSVLDLNAFERQNKAEGLGMVTEEGSGEKV
SGKDVIDEGGGRIPSKAIQVAKQVFNTLTE QGPCTGNQQSLAHSRLWDAVQFLHVRAH QMKLSQDSSQIELLKELMDLQKDMVVMLL MLEGNVVNGTIGKQMVPMLVESSNNVEMI KFFDMFLKLKDLTSSDTTKEYPPDGKGVIS RDFHKAMESHKHVTQSETEFLLSCAETDEN TLDYEEFVKRFHERAKDIGNAVLITINLSS MPNDTRLQTFLELAGSVLNYFQPFLGREIM SAKRIERVYFEISESSRTQWEKPQVKESKRQ FDVVNEGGEKEKMELFVNFCEDTIFEMQLA QISESDLNERSANKEESEREPEGQFRMAP SILTVRSALFALRYNLTLMRMLSLKSLKAQ KKVKKMTVKDMVTAFFSSYWSIFMTLLHFI ASVFRGFFRIICSLLGGSLVEGAKKIKVAEL ANMPDFTQDEVRGDGEGEREKJKEE LTDLKELTESDLLSDIEGDLKREEGQGYKL HNPNAGLSDLMSNPVPMPEVQEKPGGKAE EEEKEEKEETSSPEKAEGGEKEEKAKE KGKOKLRQLHTHRYGEPEVPESAFWKIMA QQKLLMYFARNFYNMRMLAFVAANHFIL FYKVSTSSVVGGKELPTRSSSENAKVTSLDS SHRIIAVHYVLEESSGYMEPTVRILPIHTVIS FCIGYYCLKVFLVIFKREEKVARKUEFDGL) TEQPSEDDIKGQWDRLVINTQSFPNNVNDK KLGVVFIDNSFLYLAWYMT FSDAREKKKPKXDSSLSAVLNSDKYQMW KLGVVFIDNSFLYLAWYMT 553 1903 A 4199 31 767 LPELNGRGRGGRERARGARGSAA AYGLYLRIDKGRLQCLNESSEGGGRERTQQVA LPELNGRGRGGGCCRAARGAAGSAA AYGLYLRIDKGRLQCLNESSEGSGRGVFEPP ERADDRSKFVESDADEELLFNIPFIGHVKL GIIMGEDDDSHFSSMRLYKNIPGMSFDDTEI EPDQTTSLNRDLTGELEYATKINSGNSVYHL HISKNFGADTKVFYIGRGWTELRRETV CNYEASANPADHRVHQVTPQTHFIS EPDQTTSLNRDLTGELEYATKINSGNSTDTG GIPCTEMGNPDNANVTGEFFAHIYCFKTFIS EICIKACKNLAYGEEKKKCNPYVYKTLPL RSSQGKRKTGVORNTVDPTFQETLXYOVAQ QLVTRQLQVSVWHLGTLARRVFLGEVIPLA WDFEDSTTQSFR WHELRAKADK YSDSVPQS NGELTVRAKLVLPSRTKKLQEAQEGTDQPS HGQLCLVVLGAKNLFVRPGGTLNSFVKGCL LPDQQKLRLKSPVLRKQAQCTBURSFSFG TAQLRQSSLELTVWDQALFGMNDRLLGGT RLGSKGDTAVGGONCACORCHWSFYFSG TAQLRQSSLELTVWDQALFGMNDRLLGGT RLGSKGDTAVGGONCACORCHWSFYFSG TAQLRQSSLELTVWDQALFGMNDRLLGGT RLGSKGDTAVGGONCACORCHWSFYFSG TAQLRQSSLELTVWDQALFGMNDRLLGGT RLGSKGDTAVGGONCACORCHWSFYFSG TAQLRQSSLELTVWDQALFGMNDRLLGGT RLGSKGDTAVGGONCACORCHWSFYFSG TAQLRQSSLELTVWDQALFGMNDRLLGGT RLGSKGDTAVGROVCAUSSR LWTDMTLVLH KENKKARATREMOGSRSSEDGGSEETTPQDH NHHENERRWQQERLHREEAYVQFINELNDE DYRLMKUHNLLGTTGGETTSEELQQKLIGVKI QLASOPLLRDGTNYRDSSVPQSSNEDSLLE UNDAGTTORTSTEELQCRUCOVCAUSSR	1			ļ]		LQDDEFICDLFRFLQLLCEGHNSDFQNYLRT
GGCTGNQQSLAHSRLWDAVOGFLHVFAH QMKLSQDSSQIELLKELMDAVGKDMYVMLL MLEGNVVNGTIGKQMVDMLVESSNNVEMI KFFDMFLKLKDLTSSDTFKEYDPDGKGVISI RDFHKAMESHKHYTQSETEFLLSCAETDEN TLDYBEFVKRFHEFAKDIGFNVAVLLTIKLSE MPNDTRLQTFLELABSVLTOPFLGREMEN SAKRIERVYFEISESSRTQWEKPQVKESKQ FDVVNEGGEKEMELFVOPFLGREMEN SAKRIERVYFEISESSRTQWEKPQVKESKQ FDVVNEGGEKEMELFVOPFLGREMEN SAKRIERVYFEISESSRTQWEKPQVKESKQ FDVVNEGGEKEMELFVOPFLGREMEN SAKRIERVYFEISESSRTQWEKPQVKESKQ FDVVNEGGEKEMELFVOPFLGREMEN SAKRIERVYFEISESSRTQWEKPQVKESKQ FDVVNEGGEKEMELFVOPFLGREMEN SAKRIERVYFEISESSRTQWEKPQVKESKQ KKVKKMTVKDMVTAFFSSVWSIFMTLHFF ASVFRGFFRIIGSLLLGGSLVEGAKKIKVAEL ANMPDFTQDEVRODGEGERFLEAALPE LTDLKELTEESDLLSDIFGLDLKREGQOYKL HNPNAGLSDLMSNPVPMERPLEAALPE LTDLKELTEESDLLSDIFGLDLKREGQOYKL HNPNAGLSDLMSNPVPMERPLEAALPE LTDLKELTEESDLLSDIFGLDLKREGQOYKL HNPNAGLSDLMSNPVPMERPLEAALPE LTDLKELTEESDLLSDIFGLDLKREGGOYKL FYKVSTSSVVEGKELPTRSSSENAKVTSLDS SHRIAVHYVLEESSGYMEPPVRILPILHTUS FCIGYYCLKVPLVIFKREKEVARKLEPFGLI TEQPSEDDIEGQWDRLVHTOSFPPNY PUNK VKRKVAMDKYGEFVGROMENELLGHOMAATL FSDSARKKKPKDSSLSAVISDWSYQMW KLGVVFTDNSFLYLAWYMT LFENSGRAGRERFAFSERGGGAERTQQVA FSDSARKKKPKDSSLSAVISDWSYQMW KLGVVFTDNSFLYLAWYMT LFENSGRAGGGCCAARRYGAARGSAA AYGLYLRIDKGRLQCLNESSEGSGGVFKP ERADDRSKYKKKDSSLSAVINDFTGGHTKL GIIMGEDDDSHPSEMRI YKNIPQMSFDDTEI EPDQFTSLNRDLTGELEGYARKISPESDVTH- HISKNFGADTTKVFYIGLRGEWTELRHEVT CNNEASANFADHRVHQVTTQTHFIS EPDQFTSLNRDLTGELEGYARKISPESNVYLL HISKNFGADTTKVFYIGLRGEWTELRHEVT CNNEASANFADHRVHQVTTQTHFIS EICIKACKNLAYGEEKKKKCNPVYKYLLPI RSSQGKRKTGVORNTVDTFTGFTLKYQVA QLVTRQLQVSVWHLGTLARRVFLGEVTIPLA WPDFEDSTTGSFRWHPLAKANKYEDSVPQD NGELTVRAKLVLPSRTRKLQEAQEGTDQFSI HQLCLVVLGAKNLPVRFGGTLNSFYKGCL LPDQKKLRKSPVLRKQACQPWKHSSVFSG TAQLRQSSLELTVWDQALFGMNDRLIGGT RLGSKGDTAVGGANCACOPWKHSSVFSG TAQLRQSSLELTVWDQALFGMNDRLIGGT RLGSKGDTAVGGANCACOPWKHSSVFSG TAQLRQSSLELTVWDQALFGMNDRLIGGT RLGSKGDTAVGGANCACOPWKHSSVFSG TAQLRQSSLELTVWDQALFGMNDRLIGGT RLGSKGDTAVGGANCACOPWCHSSFFSG TAQLRQSSLELTVWDQALFGMNDRLIGGT RLGSKGDTAVGTORTWTADSEVPYRGSSTEDSLLE QLASQPALRDGATNYRGSSVFRCSSHEDSLLE UNTDMTLVLII CAMPACATOR TORTWTAND TORTWTADSEVPRESSHEDSLE UNTDMTLVLII			,					CCKDAIDEOCODNIESA Y IOA Y CAENEL TENT
OMKLSQDSSQIELLKEMDLQKDMYVMLL MLEGNVVNGTIGKQMYDMLVESSNNYEM KFFDMFLKLKDI.TSSDTFKEYDPDGKGVIS RDFHKAMESHKHYTQSETEFLLSCAETDEN TLDYEEFYKRFHEPAKDIGFNVAVLLTNLSS MPNDTRLQTFLELAESVI.NYFQPFLGREIM SAKRIERVYFEISESSRTQWEKPQVKESKRYG FDVVNEGGEKEKMELFVNFCEDTIFMQLA QISESDLNERSANKEESEKERPEEQGFRMAP SILTVRSALFALRYNILTI.MRMI.SLKSLKQ KKVKKMTVKDMYTAFFSVIFFMTLHFT ASVFRGFFRIICSLLLGGSI.VEGAKKIKVAEL ANMPDPTQDEVRGDGEGERKYSIFMTLHFT ASVFRGFFRIICSLLLGGSI.VEGAKKIKVAEL ANMPDPTQDEVRGDGEGERKYSIFMTLHFT ASVFRGFFRIICSLLLGGSI.VEGAKKIKVAEL ANMPDPTQDEVRGDGEGERKEKKER EERKEKEETESFEFKAEGGGEKEEKAKER KGKQKLRQLHTHRYGEPEVPESAFWKIIAT QQKLLNYFARNFYNMRMLATVAFAINFILL FYKVSTSSVVEGKELPTRSSSENAKVTSLDS SHRIIANHYVLEESSGYMFRVILPHATVSIF FYKVSTSSVVEGKELPTRSSSENAKVTSLDS SHRIIANHYVLEESSGYMFRVILPHATVSIF FCIGYYCLKVPLVIFKREKEVARKLEFDGL.TEQPSEDDIKGQWDRLVINTQSFFNNYWDK VKRKVAMDKYGGFVGGRFRIEHLEMDHXAAT FSDAREKKPKRDSSLSAALNSDVYOMW KLGVYFIDNSFLYLAWYMDV KUGVYFIDNSFLYLAW KUGVYFIDNSFLYLAWYMDV KUGVYFIDNSFLYLAWYMDV KUGVYFILAWYMDV KUGVYFILAWYMDV KUGVYFINSTNAM	1			ľ	1		ļ	
MLEGNVYNGTIGKOMYDMLVESSNNYEM KFFDMFLKILKDLTSSDTFEPDGRGVIS RDFHKAMESHKHYTQSETEFLLSCAETDEN TLDYBEFVKRFHEPAKDIGFNYAVLITNLSS MPNDTRLQTFLELAESVLNYTQFFLGREIM SAKRIERYYFEISESSRTQWEKPQVESKRQ FDVVNBGGEKEMBLFVNFCEDTIFEMQLA QISESDLNERSANKESEKERPEQGPRMAP SILTYRSALFALRYNILTIMSLSKSLKKQ KKVKKMTVKDMYTAFFSSYWSIPMTLLHF ASVFRGFFRIICSLLIGGSLVEGAKKIKVAEL ANMPDFLQDE VRGDGEGEGRERPLEAALPSE LTDLKELTEESDLLSDIFGLDLKREGQYKL HNPNAGLSDLMSNPYMPRVQEKFQEQKAR EEKEEKEETKSPFEAAEGGOGKERAKEE KGKOKLRQLHTHEYGEPESAPWKKHAL QOKLLNYFARNFYNMRMLALFVAFANFILL FYKVSTSSVEGKELPTRSSSMAVTSLDS SHRIAVHYVLEBSSGYMEPTVRILPILHTVIS FCIGYYCLKVPLVIFKREKEVARKLEFPGIN TEQPSEDDISGQWDRLVNTQSFPNNY DVK VKRKVMDKYGEFYGRDRISELLGMDKAAT FSDAREKKPKKDSSLSAVNSDWKYQMW KLGVVFIDNSFLYLAWYMT 553 1903 A 4199 31 767 LPENGRGGGLRRAEFSERGGAERTQQVA LPENGRGGGCKCAAERVGAARGSAA AYGLYLRIDKGRLQCLNESSEGSGGVFKP ERADDRSKFVESDADEELLFNIPFTGHVKL GIIMGEDDDSHPSEMRLYKNIPQMSFDDTEI EPDQFTSLNRDLTGELELFNIPFTGHVKL GIIMGEDDDSHPSEMRLYKNIPQMSFDDTEI EPDQFTSLNRDLTGELELFANFTGHVFKHGL GIIMGEDDDSHPSEMRLYKNIPQMSFDDTEI EPDQFTSLNRDLTGELEFAHYCFKTHSI ECKACKNLAYGEBEKKKCNPVVKYLLP RSSQGKRKTGVQNNTVDTPTQETLKYQVAP QLVTRGLQVSVWHLGTLARRVFLGEVIPLA WDFEDSTTGSFRWHPLAKAKYEDSVPNOS NGELTVRAKLVLPSRTRKLQEAQEGTDQFSI HQLCLVVLGAKNLEVRDGTLINSFYKGCL LPDQKKLRKSPVLRKQACPQWKHSSVPNSG TAQLRQSSLELTVWDQALFGMNDRLIGGT RLGSKGDTAVGGACACQNWKHSSVPNSG TAQLRQSSLELTVWDQALFGMNDRLIGGT RLGSKGDTAVGGACACQNWKHSSVPNSG TAQLRQSSLELTVWDQALFGMNDRLIGGT RLGSKGDTAVGGACACQNWKHSSVPNSG TAQLRQSSLELTVWDQALFGMNDRLIGGT RLGSKGDTAVGGACACQNWCWLSSSW LWTDMTLVLH KENKKARNIKRMQSRSRSDGGSEETLPQDFI NHHENERRWQQERLHREEAYYQFINEINDE DYRLMRUHNLLGTTGGETTSELQQRILDOKI QLASQPLLROGTNYRDSEVPRESSHEDSLLE UNTDMTLVLH CKRKKARNIKRMQSRSRSDGGSEETLPQDFI NHHENERRWQQERLHREEAYYQFINEINDE DYRLMRUHNLLGTTGGETTSELQQRILDOKI QLASQPLLROGTNYRDSEVPRESSHEDSLLE UNTDMTLVLH CKRKKARNIKRMQSRSRSDGGSEETLPQDFI NHHENERRWQQERLHREEAYYQFINEINDE DYRLMRUHNLLGTTGGETTSELQQRILDOKI QLASQPLLROGTNYRDSEVPRESSHEDSLLE UNDDMTLVLDT								
KFPDMFLKLKOLTSSDTFKEYPPDGKGVISS RDPHKAMESHKHYTOSFFELLSCAETDEN TLDYEEFVKRFHEPAKDIGFNVAVLLTINLSE MPNDTRLQTFLELAESVLNYFQFFLGRIEIM SAKRIERVYFEISSSTGVMEKPQVKESKRQ FDVVNEGGEKEKMELFVNFCCDTIFEMQLA QISESDLNERSANKEESKERPEEQGPRMAP SILTVRSALFALRYNILTLMRMLSLKSLKKO KKVKKMTVKDMVTAFFSSYWSIFMTLHFP ASVPRGFFRICSLLLIGGSLVEGAKKIKVAEL ANMPDPTQDEVRGDGEGGERFVILLAHPSE LTDLKELTEESDLLSDIFGLDLKREGGGYKL HNPNAGLSDLMSNPVPMPEVQEKFQEQKAA EEEKEEFEETKSEPEK AEDGEKEEKAKE KGKQKLRQLHTHRYGEPEVPESAFWKKIIA QQKLLNYFARNFYNMRMLALFVAFANFILL FYKVSTSSVVGCKELFYSENDEAVTSLDS SHRIIAVHYVLEESSGYMEPTVRILPILHTVIS FCIIGYYFCLKVPLVIFKEEVARKLEFDGL) TEQPSEDDIKGGWDRLVINTOSFPNNY WDK VKRKVMDKYVGEFYGRORDHELLGMDKAAT. FSDAREKKRKRDSSLSAVLNSDVKYQMW KLGVVFTDNSFLVAWYMT FSDAREKKRKRKDSSLSAVLNSDVKYQMW KLGVVFTDNSFLVAWYMT DPENORGAGLRAEPSERGGGAERTQQVA LPLSHGHSHGGGGCRCAAERVGAARGSAA AYGLYLRIDKGRIJ-CQLARGSAAGSAA AYGLYLRIDKGRIJ-CQLARGSAAGSAA AYGLYLRIDKGRIJ-CQLARGSAAGSAA AYGLYLRIDKGRIJ-CQLARGSAAGSAA AYGLYLRIDKGRIJ-CQLARGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG				}				I
RDFHKAMESHKHYTQSETEFLLSCAETDEN TLDVEEFVKRPHEPAKDIGPNVAVLITNLSS MPNDTRQTFLELAESVLNYFQPFLGREIM SAKKIERVYFEISESSRTQWEKPVKESKRQ FDVVNEGGEKEKMELFVNFCCEDTIFEMQLA QISESDLNERSANKEESEKEPPEQGPRMAP SILTVRSALFALRYNHI TLMRMLSLKSLKKQ KKVKKMTVKDMVTAFFSYWSIFMTLHFF ASVFRGFFRICSLLLGGSLVEGAKKIVAEL ANMPDPTQDEVRGDGEGEGERFYLEAALPSE LTDLKELTEESDLLSDIFGLDLKREGGQYKL HNPNAGLSDLMSNPVPEVOEKFQEGKAF EEEKEKEETKSEPEKAEGEDGEKEEKAKEI KGKQRLRQLHTHRYGEPSVPESAFWKILIA QQKLLNYFARNFYYMRMLALFVAFANFILL FYKVSTSSVVEGKELFTRSSSENAKVTSLDS SHRIIAVHVLEESSGYMEVPVRIPLHTVIS PCIIGYYCLKVPLVIFKREKEVARKLEFDGL) TEQPSEDDIRGQWDR.UNITQSFPNNYWDK VARKYMDKYGEFYGRPRISELLGMMDKAAL FSDAREKKPKKDSSLSAVLNSIDVKYQMW KLGVVFTIDNSFLVLAWMT 553 1903 A 4199 31 767 LPELNGRGGGRRAEPSERGGGRETQQVA LPLSHGHSIGGGCRCAAERVGAARGSAA AYGLYLRIDKGGGCCRCAAERVGAARGSAA AYGLYLRIDKGGGCCRCAAERVGAARGSAA AYGLYLRIDKGRQCCLNESREGSGRGVFKP ERADDRSSKFVESDADEELLFNIPFTGHVKL GIIMGEDDSHPSEMKVNIPOMSFDDTE EPDOTFSLNRDLTGELEYATKISRFSNVYHL HISKNFGADTTKVFYIGLREEFLRHEVT CNYEASANPADHRVHQVTPQTHFIS 554 1904 A 4200 1 961 GIPCTEMGNEDNANVTGEIEFAHTYCFKTHS) WDFEDSTTIGSFRWHPLYCFKTHS) UNFEDSTTIGSFRWHPLYCFKTHS) WDFEDSTTIGSFRWHPLYCFKTHS) HISKNFGADTTKVFYIGLREEFURDEV MOFFEDSTTIGSFRWHPLYCFKTHS HISKNFGADTTKVFYIGLREEFAHYCFKTHS GICKACKNLAYGEEKKCNPYVKTYLLP RSSGGRKKTGVQRNTVDPTFQETLXYOWA QLVTRQLQVSVWHLGTLARRVFYCKTHS HISKNFGADTTKKVPIGLREPHYCFKTHS HISKNFGADTTKKVPIGLREPHYCFKTHS HISKNFGADTTKKVPIGLREPHYCFKTHS HISKNFGADTTKKVPIGLREPHYCFKTHS HISKNFGADTTKVPYIGLREPHYCFKTHS HISKNFGADTTKVPYIGLREPHYCFKTHS HISKNFGADTTKVPYIGLREPHYCFKTHS HISKNFGADTTKVPYIGLREPHYCFKTHS HISKNFGADTTKVPYIGLREPHYCFKTHS HISKNFGADTTKVPYIGLREPHYCFKTHS HISKNFGADTTKVPYIGLREPHYCFKTHS HISKNFGADTTKVPYIGLREPHYCFKTHS HISKNFGADTTKVPYIGLREPHYCFKTHS HISKNFGADTTKVPYIGLREPHYCFKTHS HISKNFGADTTKVPYIGLREPHYCFKTHS HIGKLACKNLAYGEKKCNPYVKTYLLP RSSGGRKKTGVQRNTVPDFTGETLXYOWA QLVTRQLQVSVWHLGTLARRVFYCKTHS HIGKLACKNLAYGEKKKCNPYVKTYLLP RSSGGRKKTGVQRNTVPDFTGETLXYOWA QLVTRQLQVSVWHLGTLARRVFYCKTHS HIGKLACKNLAYGEKKCNPYVKTYLLP RSSGGRKKTGVQRNTVPDFTGEFTLYQCHTHS HIGKLACKNLAYGEKSERGGGGETLPODH HHERERVQCERLGRREF HERERYQCE					}			
TLDVEEFVKRHEPAKDIGFNVAVLITNLES MPMOTRIC OFFICIALS UNYFOOPE GREEM SAKRIER VYFEISESSRTQWEKPQVKESKRQ FDVVNEGGEKEKMELFVNFCGEDIFEMQUA QISESDLNERSANKESEKERPEQOFRMAF SILTVRSALFALRYNILTLMRMLSKSLKKQ KKVKKMTVKDMVTAPSYSWSIFMTLHFF ASVFRGFFRICSLLIGGSLVEGAKKIKVAEL ANMPDFTQDEVRGDGEEGERKPLEAALPSE LTDLKELTEESDLLSDIFGLDLKREGGQYKL HNPPNAGLSDLMSNPVPMPEVQEKKJQCKAA EEEKEEKEEKETKSEPEKAGEDGEKEEKAKEI KGKQKLRQLHTHRYGEPSVPESAFWKKILAI QOKLLNYARANFYNMALALFVAFANFILL FYKVSTSSVVEGKELPTRSSSENAKVTSLDS SHRIIAVHYVLEESSGYMEPTVRIPLHHTVIE FCIIGYYCLKVPLVIFKREKEVARKLEFDGL) TEQPSEDDIKGGWDRLVINTOSPFNY WDK KVRKYMMBKVGEFFGRFRIFELGMPMSAAT FSDAREKKFKKKDSSLSAVLNSIDVKYQMW KLGVFTIDNSFLYLAWYMT FSDAREKKFKKKDSSLSAVLNSIDVKYQMW KLGVFTIDNSFLYLAWYMT LYENGRGAGLRAEFSERGGGAERTQQVA LPLSHGHSHGGGGCCAAERVGAARGSAA AYGLYLRIDKGRLQCLNESREGSGRGVFKP ERADIDRSKFVESDADEELLFNIPFTGHVKLL GIIMGEDDDSHPSEMRLYKNIPQMSFDDTEI EPPQTFSLNRDLTGELEV ATKISRFSNVYHL HISKNFGADTTKVFYIGLRGEWTFLRRHEVY CNYEASANPADHRVHOTOTHFIS 554 1904 A 4200 1 961 GIPCTEMGRIPDNANVTGEIEFAHYCENTRIS FSSQKRKTGVQNTVDPTFQETLKYQVAP QLVTRQLQVSVWHLGTLARRVFLGEVIPLA WDFEDSTTGSFRWHIPCYCTTLSP RSSQGKRKTGVQNTVDPTFQETLKYQVAP QLVTRQLQVSVWHLGTLARRVFLGEVIPLA WDFEDSTTGSFRWHIPCHTRISP FRSQGKRKTGVQNTVDPTFQETLKYQVAP QLVTRQLQVSVWHLGTLARRVFLGEVIPLA WDFEDSTTGSFRWHIPCHTRISP FRSQGKRKTGVQNTVDPTFQETLKYQVAP QLVTRQLQVSVWHLGTLARRVFLGEVIPLA WDFEDSTTGSFRWHIPCHTNSFVKGCL LPDQQKLRLKSPVLRKQACPQWKHSFVFSG TPAQLRQSSLELTVWDQALFGMMDRLLGGT RLGSKGDTAVGGDACSSKLQWGKVLSSN LWTDMTLVLH KENKKARNINMQSRSRDGGSEETLPODH NHHENERRWQQERLHREEAYYQFINELNDE DYRLMRDHILLGTFGETTISELQRLDCKUK QLASSQPDLRDGTTNYRSVPRSSHEDSLE					1			
MPNDTRLOTTLELABSYLNYPOPFLGRIEM SAKRIERY YFEISESSRTOWEKPOVKESKRO FDVVNEGGEKEKMELFVNFCEDTIFEMQLA QISESDLNERSANKEESEKERPEQGPRMAF SILTVRSALFALR YNILTLMRMISLSLKKQ KKVKKMTVKDMVTAFFSSYWSIPMTLHFP ASVFRGFRIIGSLLIGGSLVEGAKKIKVAEL ANMPDPTQDEVRGDGEGERKPLEAALPSE LTDLKELTEESDLLSDIFGLDLKREGGGYKL HNPNAGLSDLMSNPVMPEVQEKFQEÇAA EEEKEEKETKSEPEKAGEDGEKEEKAKEI KGKQKLRQLHTHEYPEPTVEKSTAM QQKLLNYFARNFYNMRMLALFVAFANFIL FYKVSTSSVVEGKELPTRSSSENAKVTSLDS SHRIIAVHYVLEESSGYMEFTVRILPIHTVIS FCIIGYYCLKVPLVIFKREKEVARKLEFDGL) TEQPSEDDIKGQWOLIVITOSPFNNV WDK VKRKVMDKYGEFYGRDRISELLGMDKAAT FSDAREKKRPKDSSLSAVLNSIDVKYQMW VKRKVMDKYGEFYGRDRISELLGMDKAAT FSDAREKKRPKDSSLSAVLNSIDVKYQMW KLGVVYTIONSFLYLAWYMT 5533 1903 A 4199 31 767 LPELNGRGAGLRRAFPSERGGAERTQQVA LPLSHGHSHGHGGGCAAERVGAARGSAA AYGLYLRIDKGRLQCLNESREGSGRGVFKPV ERADDRSKFVESDADEELLFNPFTGHVKU GIIMGEDDSHFSENRLYKNIPQMSFDDTEI EPDQFFSLNRDLTGELEYATKISRFSNVYHL GIIMGEDDSHFSENRLYKNIPQMSFDDTEI EPDQFFSLNRDLTGELEYATKISRFSNVYHL HISKNFGADTTKVFQLGRGEWTELRHEVT CNYEASANPADHRVHQVIPQTHPIS 554 1904 A 4200 1 961 GIPCTEMGNPDNANNTGEIEFAHYCFKTHSI EICKACKNLAYGEKKKKCNPYVKTYLLPI RSSQGKRKTGVQRNTVDPIFGGETLKYQVAP QLVTRQLQVSVWHLGTLARRVFLGEVIPLA WDFEDSTTQSRF WHELRAKADKYEDSVPQS NGELTVRAKLVLPSRTRKLQBAGEGTDQPSI HGQLCLVVLGAKNLPVRPGGTLNSFVKGCL LPDQQKLRLKSPVLRKQACPQWKHSSVFSG TPAQLRQSSLELTVWDQALFGAMDRLLGGT RLGSKGDTAVGGDACSQSKLQWQKVLSSRP LWTDMTLVLH KENKKARNILKMNQSRSRSDGGSEETLPQDH NHHENERR WQQERLHREAYYQFINELDGE DYRLMRDHNLLGTPGETITSEELQQRLDGVKI QLASQPDLRDGTTNYRDSEVPRESSHEDSLLE					i			
SAKRIERVYFEISESSRTOWERPOYKEISKRO PDVVNEGGREKEKMELFVNFCEDTIFEMQLA QISESDLNERSANKEESEKERPEEQGPRMAF SILTVRSALFALRYNILTLIRMI.SLKSLKKQ KKVKKMTVKDMVTAFFSSYWSIFMTLLHF ASVFRGFFRIICSLLLGGSLVEGAKKIKVAEL ANMPDPTQDE VRGDGEEGERKPLEAALPSE LTDLKELTESDLLSDIFGLDLKREGGQYKL HNPNAGLSDLMSNPVPMPEVQEKFQEQKAH EEKEKEERKSERKAERGEGGEKEEKEAKEI KGKQKLRQLHTHRYGEPEVPESAFWKKIIA QQKLLNYFARNFYNMRMLALFVAFANFILL FYKVSTSSVVEGKELPTRSSSENAKVYSLDS SHRIIAVHYVLEESSGYMEPTVRILPILHTVIS FCIGTYCLKVPLVIFREKEKVARKLEFPGLJ TEQPSEDDIKGQWDRLVINTQSFPNNYWDK VKRKVMDKYGGFYGRDRIEBLHGMDKAAL FSDAREKKPKKDSSLSAVINSIDVKYQMW KLGVVFTDNSFLYLAWYMT 553 1903 A 4199 31 767 LPELNGRGAGIRRAEPSERGGGRETTQVA LPLSHGHSHGGGGCRCAAERVGAARGSAA AYGLYLRIDKGRLQCLNESSREGSGRGVFKPY ERADIDRSKPVESDADEELLFINPTFGHVKLI GIIMGEDDDSHPSEMRLYKNIPQMSFDDTEI EPDQTFSLNRDLTGELEYATKISRFSNYHLH HISKNFGADTTKKYYGLRGEWTELRRHEVT CNYEASANPADHRVHQVTPOTHFIS 554 1904 A 4200 I 961 GIPCTEMGNPDNANVTGEIEFAHYCFKTHSI EICKACKNLAYGEKKKKKCNPYVKTYLLPI RSSQGKRKTGVQRNTVDPTFGETLKYQVAP QLVTRQLQVSVWHIGTLARRVFLGEVIIPLA WDFEDSTTOSFRWHPLRAKADKYEDSVPQS NGGLTVVRAKLVLPSRTRKLQEAQEGTDQPS HGQLCLVVLGAKNLPYRDGTLNSFVKGCL LPDQKLRLKSPVLRQACPQWKHSFVFSG TPAQLRQSSLELTVWDQALFGMMDRLLGGT RLGSKGDTAVGGDACSGKLQWQKVLSSR LWTDMTLVLH KENKKARNILRMNQSRSRSDGGSEETLPQDH NHHENERR WQQERLHREEAYYQFINELDGD DYRLMRDHNLLGTPGETITSEELQQRLDGVK KENKKARNILRMNQSRSRSDGGSEETLPQDH NHHENERR WQQERLHREEAYYQFINELDGD DYRLMRDHNLLGTPGTETISEELQQRLDGVK	1			1				MPNDTRLOTFLELAESVLNYFOPFLGRIEIMG
FDVVNEGGEKEKMELFVNFCEDTIFEMQLÄ QISESDLNERSANKEESEKEPPEQGPRMAF SILTVRSALFALRYNILTLMRMLSLKSLKKQ KKVKKMTVKDMVTAFFSSYWSIFMTLLHF ASVFRGFFRIIGSLLLGGSLVEGAKKIKVAEL ANMPDPTQDEVRGDGEGEREKPLEAALPSE LTDLKELTEESDLLSDIFGLDLKREGQYKL HNPNAGLSDLMSNPVPMPEVQEKFQEQKAF EEEKEEKEETKSEPEKAEGEDGEKEEKAKEL KGKQKLRQLHTHSPOEPVESAFPWKKIIA' QQKLLNYFARNFYNMRMLALFVAFANFILL FYKVSTSSVVEGKELPTRSSSENAKVTSLDS SHRIIAVHYVLEESSGYMEFTVRLIPILHTVIS FCIIGYYCLKVPLVIFKEEKEVARKLEFDGL) TEQPSEDDIKGQWDRLVINTQSFPNNYWDK VKRKVAMDKYGGFYGRAPRIFELLGMMDKAAL FSDAREKKKPKKDSSLSAVLNSIDVKYQMW KLGVVFTDNSELYLAWYMT 553 1903 A 4199 31 767 LPELNGRGAGLRRAEPSERGGAERTQQVA LPLSHGHSHGGGGCRCAAERVGAARGSAA AYGLYLRIDKGRLQCLNESREGSGRGVFKPV ERADDRSKYESDADEELLFINFTGHVKUL GIIIMGEDDDSHFSERMLYKNIPQMSFDDTEI EPPQTFSLNRDLTGELEYATKISRFSNVYHL HISKNFGADTTKVPLGRGEWTELRRHEVT CNYEASANPADHRVHQVIPDTHPIS 554 1904 A 4200 1 961 GIPCTEMGNFDNANNTGEIEFAHYCFKTHSI EICKACKNLAYGEEKKKKCNPYVKTYLLPI RSSQGKRKTGVQRNTVDPTIFGTLKYQVAP QLVTRAGLQVSVWHLGTLARRVFLGEVIPLA WDFEDSTTQSFRWHPLRAKADKYEDSVPQS NGELTVRAKLVLPSSTRKLQEAQEGTDQPSI HGQLCLVVLGAKNLPVPROGTLNSFVKGCL LPDQQKLRLKSPVLRQACPQWKHSFVFSG TPAQLRQSSLELTVWQALFGAMDRLLGGT RLGSKGDTAVGGDACSQSKLQWQKVLSSRP LWTDMTLVLH KENKKARNILRMNQSRSRSDGGSEETLPQDH NHHENERR WQQERLHREEAYYQFINELDGD DYRLMRDHNLLGTTPGETTISEELQQRLDGVK VENDMTLLCH KENKKARNILRMNQSRSRSDGGSEETLPQDH NHHENERR WQQERLHREEAYYQFINELDGD VRLMRDHNLLGTTPGETTISEELQQRLDGVK VENDMTLLCH VENDMTLLCH VENDMTLLCH VENDMTLLCH VENDMTLLCH VENDMTLLCH VENDMTLLCH VENDMTLLCH VENDMTLLCH VENDMTLLCH VENDMTLLCH VENDMTLLCH VENDMTLLCH VENDMTLLCH VENDMTLLCH VENDMTLLCH VENDMTLLCH VENDMTLLCH VENDMTLLCH VENDMTLLCTPGETTSEELQQRLDGVK VLSSRP LWTDMTLLCH VENDMTLLCTPGETTISEELQQRLDGVK								SAKRIERVYFEISESSRTQWEKPQVKESKROFI
SILTVRSALFALRYNILTLMRMLSLKSLKKQ KKVKKMTVKDMVTAFFSSYWSIFMTLLHFY ASVFRGFRRICSLLGGSLVEGAKKKVARL ANMPDPTODEVRGDGEGERKPLEAALPSE LTDLKBLTEESDLSDIFGIDLKREGGQYKL HNPNAGLSDLMSNPVPMPEVQEKFQEQKAR EEEKEEKEETKSEPEKAGEGGEKEEKAKEI KGKQKLRQLHTRYGGFEVPESAFWKKIIA- QQKLLNYFARNFYNMEMLALFVAFAINFILI FYKVSTSSVUEGKELPTRYSSENAKVISLDS SHRIIAVHYVLEESSGYMEPTVRILPILHTVIS FCIIGYYCLKVPLVIFKREKEVARKLEFDGL) TEQPSEDDIKGQWDRLVINTQSFPNNY WDK VKRKVMDKYGEFYGRARBISELLGMDKAAT FSDAREKKKFKDSSLSAVINSIDVKYQMW KLGVVFTDNSFLYLAWYMT 553 1903 A 4199 31 767 LPELNGRGAGLRRAEPSERGGGAERTQQVA LPLSHGHSHGGGCCCAAAER/VGAARGSAA AYGLYLRIDKGRCCLNESSEGSGGVFKFY ERADDRSKFVESDADEELLFNIPFTGHVKLI GIIIMGEDDSHPSEMRLYKNIPQMSFDDTEI EPDQTFSLNRDLTGELEYATKISRRSNYYHL HISKNFGADTTKVFYIGLRGEWTELRRHEVT CNYEASANPADHRYVHQVTPQTHFIS 554 1904 A 4200 1 961 GIPCTEMGNFDNANVTGEIEFAHTVCFKTHSI EICIKACKNLAYGEEKKKKCNPYVKTYLLPI RSSQGKRKTGVANTVDFTFOETILKYQVAP QLVTRQLQVSVWHLGTLARRVFLGEVIPLA WDFEDSTTQSFRWHPLRAKADKYEDSVPQS NGELTVRAKU-LYSTRIKLQEAQEGTDOPSI HGQLCLVVLGAKNLPVRPDGTLNSFVKGCL LPDQKLRLKSPVLRKQACPQWKHSFVFSG TPAQLRQSSLELTVWDQALFGMNDRLLGGT RLGSKGDTVARKL-LYSTRIKLQEAQEGTDOPSI HGQLCLVVLGAKNLPVRPDGTLNSFVKGCL LPDQCKLRLKSPVLRKQACPQWKHSFVFSG TPAQLRQSSLELTVWDQALFGMNDRLLGGT RLGSKGDTAVGGDACSQSKLQWQKVLSSR LWTDMTLVLH 555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQQERLHREBAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKI QLASQPDLRDGTTNYNDSEVPRESSHEDSLLE	ĺ							FDVVNEGGEKEKMELFVNFCEDTIFEMQLAA
KKVKKMTVADMYTAFFSSYWSIFMTILHFT ASVFRGFFRICSLLLGGSLVEGAKKIKVAEL ANMPDPTQDEVRGDGEGEGERKPLEAALPSE LTDLKELTEESDLLSDIFGLDLKREGGQYKL HNPNAGLSDLMSNPVPMPEVQEKFQEQKAF EEEKEEKEETSSPEKAEGEDGEKEEKAKEI KGKQKLRQLHTHRYGEPEVPESAFWKKIIA* QQKLLNYFARNFYMREMLALFVAFANFILI FYKVSTSSVVEGKELPTRSSSSNAKVTSLDS SHRIIAVHYVLEESSGYMBPTVRILPILHTVIS FCIIGYYCLKVPLVIFKREKEVARKLEFDGI, TEQPSEDDIKGQWDRLVINTQSFPNNYWDK VKRKVAMDKYGEFYGERSHEIELLGMDKAAL* FSDAREKKFRKDSSLSAVLNSDVKYQMW KLGVVFTDNSFLYLAWYMT 553 1903 A 4199 31 767 LPELNGRGAGLRRAEFSERGGGAERTQQVA LPLSHGHSHGGGCCCAAERVGAARGSAA AYGLYLRIDKGRLQCLNESREGSGRGVFKPV ERADDRSKFVESDADEELLFNIPFTGHVKL. GIIIMGEDDDSHPSEMRLYKNIPQMSFDDTEI EPPQTFSLNRDLTGELEYAFKISRFSNVYHL! HISKNFGADTTKVFYJGLRGEWTELRHEVT CNYEASANPADHRVHQVTPQTHFIS 554 1904 A 4200 1 961 GIPCTEMGNFDNANVTGEIEFAHYCFKTHSI EICIKACKNLAYGEKKKCNPVYKTYLLPI RSSQGKRKTGVQKNTVDPTFGETLKYQVAP QLVTRQLQVSVWHLGTLARRVFLGEVIIPLA WDFEDSTTQSFRWHPLRAKADKYEDSVPQS NGELTVRAKLVLPSRTRKLQEAGEGTDQPSI HGQLCLVVLGAKNLPVRPGGTLNSFVKGCL LPDQKLRLKSPVLRKQACPQWKHSFVFSG TPAQLRQSSLELTVWDQALFGMNDRLLGGT RLGSKGDTAVGGDACSQSKLQWQKVLSSPR LWTDMTLVH 555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQQERLJREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKI QLASQPDLRDGTTNYDSEVPRESSHEDSLLE DYRLMRDHNLLGTPGEITSEELQCRLDGVKI								QISESDLNERSANKEESEKERPEEQGPRMAFF
ASVPRGFRRICSLLLGGSLVEGAKKIKVABL ANMPDPTQDE VPGDGEGERKPLEAALPSE LTDLKELTEESDLLSDIFGLDLKREGGQYKL HNPNAGLSDLMSNPVPMPEVQEKFQEQKAF EEKEKEKEETSEPKAFGEGGEKEEKAKEI GEKEKEKETSEPKAFGEGGEKEEKAKEI GQKLLNYFARNFYNMRMLALFVAFANFILL FYKVSTSSVVEGKELPTRSSSENAKVTSLDS SHRIIAVHYVLEESSGYMEPTVRILPILHTVIS FCIIGYYCLKVPLVIFKREKEVARKLEPGGL) TEQPSEDDIRGQMDRLVINTOSFPNNYWDK VKRKWMDKVGEFYGERBRISELLGMDKAAL FSDAREKKKPKKDSSLSAVLNSIDVYQMW KLGVVFTDNSFLYLAWYMT 553 1903 A 4199 31 767 LPELNGRGAGLRRAEPSERGGGARTQQVA LPLSHGHSHGGGGCRCAAERVGAARGSAA AYGLYLRIDKGRLQCLNESREGSGRGVFKP ERADIDRSKFVESDADEELLFNIPFTGHVKL. GIIMGEDDDSHPSEMRLYKNIPQMSFDDTEI EPPQTFSLNRDLTGELEYATKISRFSNVYHL. HISKNFGADTTKVFYIGLRGEWTELRRHEVT CNYEASANPADHRVHQVTPQTHFIS 554 1904 A 4200 1 961 GIPCTEMGNFDNANVTGEIEFAHYCFKTHSI EICIKACKNLAYGEEKKKCNPYVKTVLLPI RSSQGKRKTGVQRNTVDFTFQTLKYQVAP QLVTRQLQVSVWHLGTLARRVFLGEVIPLA WDFEDSTTQSFRWHPLRAKADK YEDSVPQS NGELTVRAKLVLPSRTRKLQEAQEGTDQPSI HGQLCLVVLGAKNLPVRPDGTLNSFVKGCL LPDQKLRLKSPVLRKQACPQWKHSFVFSG TRAQLRQSSLELTVWDQALFGMNDRLLGGT RLGSKGDTAVGGGDACSQSKLQWQKVLSSR LWTDMTLVLH 555 1905 A 4211 331 2419 KENKKARPLRMQSRSRSDGGSEETLPQDH NHEINERRWQQERLHREAYYQPINELNDE DYRLMRDHNLLGTPGETISEELQQRLDGVKI QLASQPDLRDGTYRRDSEVPRESSHEDSLE								SILTVRSALFALRYNILTLMRMLSLKSLKKQM
ANMPDPTQDEVRGDGEGERKPLEAAL PSE LTDLKELTEESDLLSDIFGLDLKREGGGYKLL HNPNAGLSDLMSNPVPMPEVQEKFGEQEXAH EEBKEEKEETK SEPEK AEGEDGEKEEKAKEI KGKQKLRQLHTHRYGEPEVPESAFWKKIIA' QQKLLNYFARNFYNMRMLALFVAFANFILI FYKVSTSSVVEGKELPTRSSSENAKVTSLDS SHRIIAVHYVLEESSGYMEPTVRILPILHTVIS FCIIGYYCLKVPLVIFKREKEVARKLEFDGL) TEOPSEDDIKGQWDRLVINTQSFPNNYWDK VKRKVMDKYGEFYGRDRISELLGMDKAAL' FSDAREKKPKDSSLSAVLNSIDVKYQMW KLGVVFTDNSFLYLAWYMT LPELNGRGAGLRRAEPSERGGGAERTQQVA LPLSHGHSHGGGGCRCAAERVGAARGSAA AYGLYLRDKGRLQCLNESREGSGRGVFKP' ERADDRSKFVESDADEELLFNIPFTGHVKLL GIIIMGEDDDSHPSEMRLYKNIPQMSFDDTEI EPPQTFSLNRDLTGELEVATKISRFSNYYHL. HISKNFGADTIKVFYIGLREWTELRRHEVT CNYEASANPADHRVHQVTPQTHFIS 554 1904 A 4200 1 961 GIPCTEMGNFDNANVTGEIEFAIHVCFKTHSI EICIKACKNLAYGEEKKKCNPYVKTYLLPI RSSQGKRKTGVQRNTVDPTTQETLKYQVAP QLVTRQLQVSVWHLGTLARRVFLGEVIPLA WDFEDSTTQSFRWPLRAKADKYEDSVPQS NGELTVRAKLVLPSRTRKLQEAQEGTDQPSI HGQLCLVVLGAKNLPVRPDGTLNSFVKGCL LPPQQKLRLKSPVLRKQACPQWKHSFVFSG TPAQLRQSSLELTVWDQALFGMNDRLLGGT RLGSKGGTAVGGDACSQSKLQWQKVLSSRN LWTDMTLVLH 555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQQERLHREEAYYQPINELNDE DYRLMRDHNLLGTTRGESSEDLGURGVKI QLASQPDLRDGTNYRDSEVPRESSHEDSLE	ı				1			KKVKKMTVKDMVTAFFSSYWSIFMTLLHFV
LTDLKELTĒESDLLSDIFGLDLKREGGQYKL HNPNAGLSDLMSNPVPMPEVQEKFQEQKAR EEEKEKEETKSEPEKAGEGGEKEEKAKEI KGKQKLRQLHTHRYGEPEVPESAFWKKIIA' QQKLLNYFARNFYNMRMLALFVAFANNFILD FYRVSTSSVVEGKELPTRSSSENAKVTSLDS SHRIIAVHYVLEESSGYMEPTVRILPILHTVIS FCIIGYYCLKVPLVIFKREKEVARKLEPOGL) TEQPSEDDIKGQMDRLVINTOSPPNNY WDK VKRKVMDKYGEFYGRDRISELLGMDKAAL FSDAREKKKPKDSSLSAVLNSIDVKYQMW KLGVVFTDNSFLYLAWYMT 553 1903 A 4199 31 767 LPELNGRGAGLRRAEPSERGGAERTQQVA LPLSHGHSHGGGGCRCAAERVGAARGSAA AYGLYLRIDKGRLQCLNESREGSGRGVKFV ERADIDRSKFVESDADEELLFNIPFTGHVKL. GIIIMGEDDDSFIPSEMRLYKNIPQMSFDDTEI EPDQTFSLNRDLTGELEYATKISRFSNVYHL: HISKNFGADTITKVFYIGLRGEWTELRRHEVT CNYEASANPADHRVHQVTPQTHFIS 554 1904 A 4200 1 961 GIPCTEMONIVTGEIEFAIHYCFKTHSI EICIKACKNLAYGEEKKKKCNPYVKTYLLPI RSSQGKRKTGVQRNTVDFTPQETLKYQWAP QLVTRQLQVSVWHLGTLARRVFLGEVIIPLA WDFEDSTTQSFRWHPLRAKADKYEDSVPQS NGELTVRAKL VLPSRTRKLQEAQEGTDQPSI HGQLCLVVLGAKNLPVRPDGTLNSFVKGGL LPPQQKLRLKSSPVLRKQACPQWKHSVFVSG TPAQLRQSSLELTVWDQALFGMNDRLLGGT RLGSKGDTAVGGDACSQSKLQWGKVLSSRN LWTDMTLVLH 555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQQELHREEAYYQDRILDGVKI	1					'		ASVFRGFFRIICSLLLGGSLVEGAKKIKVAELL
HNPNAGLSDLMSNPVPMPEVOEKFOEOKARE EEEKEKETKSEPEKAEGEDGEKEEKAKE KGKQKLRQHTHRYGEPEVPESAFWKKIIA' QQKLLNYFARNFYNMRMLALFVAFAINFILI FYKVSTSSVVEGKELPTRSSSENAKVISLOS SHRIIAVHYVLEESSGYMEPTVRILPILHTVIS FCIGYYCLKVPLVIFKREKEVARKLEFDGLN TEQPSEDDIKGQWDRLVINTQSFPNNYWDK VKRKVMDKYGEFYGRDRISELLGMPMAAT. FSDAREKKPFKDSSLSAVLNSIDVKYQMW KLGVVFTDNSFLYLAWYMT 553 1903 A 4199 31 767 LPELNGRGGGIRRAEPSERGGAERTQQVA LPLSHGHSHGGGGCRCAAER/VGAARGSAA AYGLYLRIDKGRLQCLNESREGSGRGGYFKPY ERADDRSKFVESDADEELLFINPFTGHYKLL GIIMGEDDDSHPSEMRLYKNIPQMSFDDTEI EPDQTFSLNRDLTGBLEYATKISRFSNVYHLH HSKNFGADTTKVPYIGLRGEWTELRRHEVT CNYEASANPADHRVHQVTPQTHFIS 554 1904 A 4200 1 961 GIPCTEMGNFDNANVTGEIEFAHYCFKTHLY KISKNFGADTTKVPYIGLRGEWTELRRHEVT CNYEASANPADHRVHQVTPQTHFIS FSDAGREKKKCNPYVKTYLLPI RSSQGKRKTGVQRNTVDPTFQETLKYQVAP QLVTRQLQVSVWHJGTLARRVFLGEVIIPLA WDFEDSTTOSFRWHPLRAKADKYEDSVPQS NGELTVRAKLVLPSRTKKLQEAQEGTDQPSI HGQLCLVVLGAKNLPVRPDGTLNSFVKGCL LPPDQGKLKLSPVLRKQACPQWKHSFVFSG TPAQLRQSSLELTVWDQALFGMNDRLLGGT RLGSKGDTAVGGDACSQSKLQWQKVLSSRP LWTDMTLVLH 555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQRLDGVKI QLASQPDLRDGTNYRDSEVPRESSHEDSLLE	ı							ANMPDPTQDEVRGDGEEGERKPLEAALPSED
EEEKEKETKSEPEKAEGEDGEKEKAKEI KGKQKLRQLHTHRYGEPEVPESAFWKKIIA QQKLLNYFARNFYNMRMLALFVAFANFILI FYKVSTSSVVEGKELPTRSSSENAKVTSLDS SHRIIAVHYVLEESSGYMEPTVRILPILHTVIS FCIIGFYYCLKVPLVIFKREKEVARKLEFDGLI TEQPSEDDIKGQWDRLVINTQSFPNNYWDK VKRKWMKYGEFYGRDRISEIL-GMDKAAL- FSDAREKKPKKDKSLSAAVLNSIDVKYQMW KLGVVFTDNSFLYLAWYMT 553 1903 A 4199 31 767 LPELNGRGAGLRRAEPSERGGGAERTQQVA LPLSHGHSHGRGGCRCAAER/VGAARGSAA AYGLYLRIDKGRLQCLNESREGSGRGVFKPY ERADIDRSKFVESDADEELLFNIPFTGHYKLL GIIIMGEDDDSHPSEMRLYKNIPQMSFDDTEI EPPDQTFSLNRDLTGELEYATKISRFSNVYHL. HISKNFGADTTKVFYIGLRGEWTELRHEVT CNYEASANPADHRVHQVTTQTHFIS 554 1904 A 4200 1 961 GIPCTEMONFDNANVTGEIEFAHYCFKTHSI EICIKACKNLAYGEEKKKKCNPYVKTYLLPI RSSQGKRKTGVQRNTVDPIFQETLKYQVAP QLVTRQLQVSVWHLGTLARRVFLGEVIIPLA WDFEDSTTQSFRWHPLRAKADKYEDSVPQS NGELTVRAKLVLPSRTRKLQEAQEGTDQPSI HGQLCLVVLGAKNLPVRPDGTLNSFVKGCL LPDQKLRLKSPVLRKQACPQWKHSFVPSG TPAQLRQSSLELTVWDQALFGMNDRLLGGT RLGSKGDTAVGGDACSQSKLQWQKVLSSRN LWTDMTLVLH 555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQQERLHREEAYYQFINELNDE DYRLMRDHVLLGTTGEITSEELQQRLDGVKI QLASQPDLRDGTNYRDSEVPRESSHEDSLLE					ĺ			
KGKQKLRQLHTHRYGEPEVPESAFWKKIIAY QQKLLNYFARNFYNMRMLALFVAFARNFILL FYKVSYSVYGKELFPTRSSSEBAKVTSLDS SHRIIAVHYVLEESSGYMEPTVRILPILHTVIS FCIIGYYCLKVPLVIFKREKEVARKLEFDGLY TEQPSEDDIKGQWDRLVINTQSFPNNYWDK VKRKVMDKYGGFYGRDRISELGMMKAAL FSDAREKKKPKKDSSLSAVLNSIDVKYQMW KLGVVFTDNSFLYLAWYMT 553 1903 A 4199 31 767 LPELNGGAGLRRAEPSERGGGAERTQQVA LPLSHGHSHGGGGCRCAAER/VGAARGSAA AYGLYLRIDKGRQCLNESREGSGGVFKPY ERADDRSKFVESDADEELLFNIPFTGHVKLI GIIIMGEDDDSHPSEMRLYKNIPQMSFDDTEI EPDQTFSLNRDLTGELEYATKISRFSNVYHL. HISKNFGADTTKVFYIGLRGEWTELRRHEVT CNYEASANPADHRVHQVTPQTHPIS 554 1904 A 4200 1 961 GIPCTEMGNFDNANVTGEIEFAIHYCFKTHSI EICIKACKNLAYGEEKKKKCNPYVKTYLLPI RSSQGKRKTGVQRNTVDPIFQETLKYQVAP QLVTRQLQVSVWHLGTLARRVFLGEVIIPLA WDFEDSTTQSFRWHPLRAKADKYEDSVPQS NGELTVRAKLVLPSRTRKLQEAQEGTDQPSI HGQLCLVVLGAKNLPVRPDGTLNSFVKGCL LPDQQKLRLKSPVLRKQACPQWKHSFVFSG TPAQLRQSSLELTVWDQALFGMNDRLLGGT RLGSKGDTAVGGDACSQSKLQWGKVLSSPR LWTDMTLVLH 555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQQRLHREEAYYQFINELNDE DYRLMRDHNLLGTGETISEELQQRLDGVXIG QLASQPDLRDGTNYRDSEVPRESSHEDSLLE	-							
QQKLLNYFARNFYNMRMLALFVAFAINFILI FYKVSTSSVVEGKELPTRSSENAKVTSLDS SHRIIAVHYVLEESSGYMEPTVRILPILHTVIS FCIIGYYCLKVPLVIFKREKEVARKLEFDGLN TEQPSEDDIKGQWDRLVINTQSFPNNYWDK VKRKVMKVGEFYGDRISELLGMDKAAL FSDAREKKKPKDSSLSAVLNSIDVKYQMW KLGVVFTDNSFLYLAWYMT 553 1903 A 4199 31 767 LPELNGRGAGGRRAEPSERGGAERTQQVA LPLSHGHSHGGGGCRCAAERVGAARGSAA AYGLYLRIDKGRLQCLNESREGSGRGVFKPV ERADIDRSKFVESDADEELLFINFFTGHVKL GIIDHORSKFVESDADEELLFINFFTGHVKL GIISHGSDDSHPSEMRLYKNIPQMSFDDTEI EPDQTFSLNRDLTGELEYATKISRFSNVYHL HISKNIFGADTTKVFYIGLRGEWTELRRHEVT CNYEASANPADHRVHQVTPQTHPIS 554 1904 A 4200 1 961 GIPCTEMGNFDNANVTGEIEFAHYCFKTHSI EICIKACKNLAYGEEKKKKCNPYVKTVLLPI RSSQGKRKTGVQRNTVDPTFQETLKYQVAP QLVTRQLQVSVWHLGTLARRVFLGEVIPLA WDFEDSTTQSFRWHPLRAKADKYEDSVPQS NGELTVRAKLVLPSRTRKLQEAQEGTDQPSI HGQLCLVUGAKNLPVRPDGTLNSFVKGCL LPDQQKLRLKSPVLRKQACPQWKHSFVPSG TPAQLRQSSLELTVWDQALFGMNDRLLGGT RLGSKGDTAVGGDACSQSKLQWQKVLSSRP LWTDMTLVLH 555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQQERLHREEAYYQFINELNDE DYRLMRDNLLGTTGETISEELQQRLDGVKI QLASQPDLRDGTNYRDSEVPRESSHEDSLLE								L
FYKVSTSSVVEGKELPTRSSSENAKVTSLDS SHRIIAVHYVLEESSGYMEPTVRIJLHTVIE FCIGYYCLKVPLVIFKREKEVARKLEFDGLJ TEQPSEDDIKGQWDRLVINTQSFPNNYWDK VKRKYMDKYGEFYGRDRISELLGMDKAAT: FSDAREKKKPKKDSSLSAVLNSIDVKYQMW KLGVVFTDNSFLYLAWYMT 553 1903 A 4199 31 767 LPELNGRGAGLRRAEPSERGGAERTQQVA LPLSHGHSHGGGGCRCAAER/VGAARGSAA AYGLYLRIDKGRLQCLNESREGSGRGVFKPV ERADIDRSKFVESDADEELLFNIPFTGHVKL. GIIIMGEDDDSHPSEMRLYKNIPQMSFDDTEI EPDQTFSLNDLTGELEYATKISRFSNYHL.S HISKNFGADTTKVPYIGLRGEWTELRRHEVT CNYEASANPADHRVHQVTPQTHPIS EICIKACKNLAYGEEKKKKCNPYVKTYLLPI BICIKACKNLAYGEEKKKKCNPYVKTYLLPI RSSQGKRTGVQRNTVDPITFQETLKYQVAP QLVTRQLQVSVWHLGTLARRVFLGEVIPLA WDFEDSTTQSFRWHPLRAKADKYEDSVPQS NGELTVRAKLVLPSRTRKLQEAQEGTDQPSI HGQLCLVVLGAKNLPVRPDGTLNSFVKGCL LPDQQKLRLKSPVLRKQACPQWKHSFVFSG TPAQLRQSSLELTVWDQALFGMNDRLLGGT RLGSKGDTAVGGDACSQSKLQWQKVLSSP LWTDMTLVH 555 1905 A 4211 331 2419 KENKKARNLRMMQSRSRSDGGSEETLPQDH NHHENERRWQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGETISEELQQRLDGVKI QLASQPDLRDGTNYRDSEVPRESSHEDSLLE								
SHRIIAVHYVLEESSGYMEPTVRILPILHTVIS FCIIGYYCLKVPLVIFKREKEVARKLEFDGI, TEQPSEDDIKGQWDRLVINTQSFPNNYWDK VKRKYMDKYGEFYGRDRISELLGMPKAAL FSDAREKKKPKDSSLSAVLNSIDVKYQMW KLGVVFTDNSFLYLAWYMT 553 1903 A 4199 31 767 LPELNGGAGLRRAEPSERGGAERTQQVA LPLSHGHSHGGGGCRCAAER/VGAARGSAA AYGLYLRIDKGRLQCLNESREGSGRGVFKPV ERADIDRSKFVESDADEELLFNIPFTGHVKL. GIIIMGDDDSHPSEMRLYKNIPQMSFDDTEI BPDQTFSLNRDLTGELEYATKISRFSNVYHL. HISKNFGADTTKVFYIGLRGEWTELRRHEVT CNYEASANPADHRVHQVTPQTHFIS 554 1904 A 4200 1 961 GIPCTEMGNFDNANVTGEIEFAHYCFKTHSI EICIKACKNLAYGEEKKKKCNPYVKTYLLPI RSSQGKRKTGVQRNTVDPIFQETLKYQVAP QLVTRQLQVSVWHLGTLARRVFLGEVIPLA WDFEDSTTQSFRWHPLRAKADKYEDSVPQS NGELTVRAKLVLPSRTRKLQEAQEGTDQPSI HGQLCLVVLGAKNLPVRPDGTLNSFVKGCL LPDQQKLRLKSPVLRKQACPGWKHSFVFSG TPAQLRQSSLELTVWDQALFGMNDRLLGGT RLGSKGDTAVGGDACSQSKLQWQKVLSSRN LWTDMTLVLH 555 1905 A 4211 331 2419 KENKKARNLRMMQSRSRSDGGSEETLPQDH NHHENERRWQQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKI QLASQPDLRDGTNYRDSEVPRESSHEDSLLE					,			
FCIIGYYCLKVPLVIFKREKEVARKLEFDGLY TEQPSEDDIKGQWDRLVINTQSFPNNYWDK VKRKVMDKYGEFYGRDRISELLGMDKAAL FSDAREKKPKNSDSLSAVLNSIDVKYQMW KLGVVFTDNSFLYLAWYMT 553 1903 A 4199 31 767 LPELNGRGAGLRRAEPSERGGAERTQQVA LPLSHGHSHGGGGCAAER/VGAARGSAA AYGLYLRIDKGRLQCLNESREGSGRGVFKPV ERADIDRSKFVESDADEELLFNIPFTG\HVKLI GIIMGEDDDSHPSEMRLYKNIPQMSFDDTEI EPDQTFSLNRDLTGELEYATKISRFSNVYHL! HISKNFGADTTKVFYIGLRGEWTELRRHEVT CNYEASANPADHRVHQVTPQTHEIS 554 1904 A 4200 1 961 GIPCTEMGNFDNANVTGEIEFAIHYCFKTHSI EICIKACKNLAYGEEKKKCNPYVKTYLLPI RSSQGKRKTGVQRNTVDPTFQETLKYQVAP QLVTRQLQVSVWHLGTLARRVFLGEVIIPLA WDFEDSTTQSFRWHPLRAKADKYEDSVPGS NGELTVRAKLVLPSRTRKLQEAQEGTDQPSI HGQLCLVVLGAKNLPVRPDGTLNSFVKGCL LPPQQKLRLKSPVLRKQACPQWKHSFVFSG TPAQLRQSSLELTVWDQALFGMNDRLLGGT RLGSKGDTAVGGDACSQSKLQWQKVLSSIN LWTDMTLVLH 555 1905 A 4211 331 2419 KENKKARNILRMNQSRSRSDGGSEETLPQDH NHHENERRWQQERLHREEAYYQFINELNDG DYRLMRDHNLLGTFGEITSEELQQRLDGVKI QLASQPDLRDGTNYRDSEVPRESSHEDSLLE	İ							1
TEQPSEDDIKGQWDRLVINTQSFPNNYWDK VKRKVMDKYGEFYGRDRISELLGMPKAAT: FSDAREKKKPKDSSLSAVLNSIDVKYQMW KLGVVFTDNSFLYLAWYMT 553 1903 A 4199 31 767 LPELNGRGAGLRRAEPSERGGAERTQQVA LPLSHGHSHGGGGCRCAAER/VGAARGSAA AYGLYLRIDKGRLQCLNESREGSGRVFKPY ERADDRSKFVESDADEELLFNIPFTGHVKLI GIIMGEDDDSHPSEMRLYKNIPQMSFDDTEI EPDQTFSLNRDLTGELEYATKISRFSNVYHL: HISKNFGADTTKVFYIGLRGEWTELRRHEVT CNYEASANPADHRVHQVTPQTHFIS 554 1904 A 4200 1 961 GIPCTEMGNFDNANVTGEIEFAIHYCFKTHSI EICIKACKNLAYGEEKKKKCNPYVKTYLLPI RSSQGKRKTGVQRNTVDPTFQETLKYQVAP QLVTRQLQVSVWHLGTLARRVFLGEVIIPLA WDFEDSTTQSFRWHPLRAKADKYEDSVPQS NGELTVRAKLVLPSRTRKLQEAQEGTDQPSI HGQLCLVVLGAKNLPYRPDGTLNSFVKGCL LPDQCKLRKSPVLRKQACPQWKHSFVFSG TPAQLRQSSLELTVWDQALFGMNDRLLGGT RLGSKGDTAVGGDACSQSKLQWGKVLSSRN LWTDMTLVLH 555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQQERLHREEAYYQFINELNDE DYRLMRDHNLLGTTGGTISEELQQRLDGVKI QLASQPDLRDGTNYRDSEVPRESSHEDSLLE					Į į			
VKRKVMDKYGEFYGRDRISELLGMDKAAL FSDAREKKPKKDSSLSAVLNSIDVKYQMW KLGVVFTDNSFLYLAWYMT S53								
FSDAREKKKPKDSSLSAVLNSIDVKYQMW KLGVVFTDNSFLYLAWYMT 553 1903 A 4199 31 767 LPELNGRAGGIRRAEPSERGGAERTQQVA LPLSHGHSHGGGGCRCAAER/VGAARGSAA AYGLYLRIDKGRLQCLNESREGSGRGVFKPV ERAD\DRSKFVESDADEELLFNIPFTG\HVKLI GIIIMGEDDDSHPSEMRLYKNIPQMSFDDTEI EPDQTFSLNRDLTGELEYATKISRFSNVYHLI HISKNFGADTTKVFYIGLRGEWTELRRHEVT CNYEASANPADHRVHQVTPQTHFIS 554 1904 A 4200 1 961 GIPCTEMGNFDNANVTGEIEFAIHYCFKTHSJ EICIKACKNLAYGEEKKKCNPYVKTYLLPI RSSQGKRKTGVQRNTVDPIFQETLKYQVAP QLVTRQLQVSVWHLGTLARRVFLGEVIIPLA WDFEDSTTQSFRWHPLRAKADKYEDSVPQS NGELTVRAKLVLPSRTRKLQEAQEGTDQPSI HGQLCLVVLGAKNLPVRPDGTLNSFVKGCL LPDQQKLRLKSPVLRKQACPQWKHSFVFSG TPPAQLRQSSLELTVWDQALFGMNDRLLGGT RLGSKGDTAVGGDACSQSKLQWGVVLSSPN LWTDMTLVH 555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKI QLASQPDLRDGTNYRDSEVPRESSHEDSLLE							·	
1903 A 4199 31 767 LPELNGRGAGLRRAEPSERGGGAERTQQVA LPLSHGHSHGGGGCRCAAER/VGAARGSAA AYGLYLRIDKGRLQCLNESREGSGRGVFKFW ERAD\DRSKFVESDADEELLFNIPFTGHVKLI GIIIMGEDDDSHPSEMRLYKNIPQMSFDDTEI EPDQTFSLNRDLTGELEYATKISRFSNVYHLI HISKNFGADTTKVFYIGLRGEWTELRRHEVT CNYEASANPADHRVHQVIPQTHFIS	+							FSDAREKKKPKKDSSLSAVLNSIDVKYOMW
LPLSHGHSHGGGGCRCAAER/VGAARGSAA AYGLYLRIDKGRLQCLNESREGSGRGVFKPV ERAD\DRSKFVESDADEELLFNIPFTG\HVKLI GIIMGEDDDSHPSEMRLYKNIPQMSFDDTEI EPPQTFSLNRDLTGELEYATKISRFSNVYHLS HISKNFGADTTKVFYIGLRGEWTELRHEVT CNYEASANPADHRVHQVTPQTHFIS 554 1904 A 4200 1 961 GIPCTEMGNFDNANVTGEIEFAIHYCFKTHS) EICIKACKNLAYGEEKKKKCNPYVKTYLLPI RSSQGKRKTGVQRNTVDPTFQETLKYQVAP QLVTRQLQVSVWHLGTLARRADKYEDSVPQS NGELTVRAKLVLPSRTRKLQEAQEGTDQPSI HGQLCLVVLGAKNLPVRPDGTLNSFVKGCL LPDQQKLRLKSPVLRKQACPQWKHSFVFSG TPAQLRQSSLELTVWDQALFGMNDRLLGGT RLGSKGDTAVGGDACSQSKLQWQKVLSSP\ LWTDMTLVLH 555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKI QLASQPDLRDGTNYRDSEVPRESSHEDSLLE								
AYGLYLRIDKGRLQCLNESREGSGRGVFKPV ERAD\DRSKFVESDADEELLFNIPFTG\HVKLI GIIIMGEDDDSHPSEMRLYKNIPQMSFDDTEI EPDQTFSLNRDLTGELEYATKISRFSNVYHLI HISKNFGADTTKVFYIGLRGEWTELRRHEVT CNYEASANPADHRVHQVIPQTHPIS 554 1904 A 4200 1 961 GIPCTEMGNFDNANVTGEIEFAIHYCFKTHSI EICIKACKNLAYGEEKKKKCNPYVKTYLLPI RSSQGKRKTGVQRNTVDPIFQETLKYQVAP QLVTRQLQVSVWHLGTLARRVFLGEVIIPLA WDFEDSTTQSFRWHPLRAKADKYEDSVPQS NGELTVRAKLVLPSRTRKLQEAQEGTDQPSI HGQLCLVVLGAKNLPVRPDGTLNSFVKGCL LPDQQKLRLKSPVLRKQACPQWKHSFVFSG TPAQLRQSSLELTVWDQALFGMNDRLLGGT RLGSKGDTAVGGDACSQSKLQWQKVLSSP\ LWTDMTLVLH 555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERR WQQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKI QLASQPDLRDGTNYRDSEVPRESSHEDSLLE		553	1903	A	4199	31	767	LPELNGRGAGLRRAEPSERGGGAERTQQVAA
ERAD\DRSKFVESDADEELLFNIPFTG\HVKLI GIIIMGEDDDSHPSEMRLYKNIPQMSFDDTEI EPDQTFSLNRDLTGELEYATKISRFSNVYHLS HISKNFGADTTKVFYIGLRGEWTELRRHEVT CNYEASANPADHRVHQVTPQTHFIS 554 1904 A 4200 1 961 GIPCTEMGNFDNANVTGEIEFAIHYCFKTHSI EICIKACKNLAYGEEKKKKCNPYVKTYLLPI RSSQGKRKTGVQRNTVDPIFQETLKYQVAP QLVTRQLQVSVWHLGTLARRVFLGEVIIPLA WDFEDSTTQSFRWHPLRAKADKYEDSVPQS NGELTVRAKLVLPSRTRKLQEAQEGTDQPSI HGQLCLVVLGAKNLPVRPDGTLNSFVKGCL LPDQQKLRLKSPVLRKQACPQWKHSFVFSG TPAQLRQSSLELTVWDQALFGMNDRLLGGT RLGSKGDTAVGGDACSQSKLQWQKVLSSP LWTDMTLVLH 555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKI QLASQPDLRDGTNYRDSEVPRESSHEDSLLE	1	ļ						LPLSHGHSHGGGGCRCAAER/VGAARGSAAC
GIIIMGEDDDSHPSEMRLYKNIPQMSFDDTEI EPDQTFSLNRDLTGELEYATKISRFSNVYHLS HISKNFGADTTKVFYIGLRGEWTELRRHEVT CNYEASANPADHRVHQVTPQTHFIS 554 1904 A 4200 I 961 GIPCTEMGNFDNANVTGEIEFAIHYCFKTHSI EICIKACKNLAYGEEKKKKCNPYVKTYLLPI RSSQGKRKTGVQRNTVDPIFQETLKYQVAP QLVTRQLQVSVWHLGTLARRVFLGEVIPLA WDFEDSTTQSFRWHPLRAKADK YEDSVPQS NGELTVRAKLVLPSRTRKLQEAQEGTDQPSI HGQLCLVVLGAKNLPVRPDGTLNSFVKGCL LPDQQKLRLKSPVLRKQACPQWKHSFVFSG TPAQLRQSSLELTVWDQALFGMNDRLLGGT RLGSKGDTAVGGDACSQSKLQWQKVLSSPN LWTDMTLVLH 555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKI QLASQPDLRDGTNYRDSEVPRESSHEDSLLE								AYGLYLRIDKGRLQCLNESREGSGRGVFKPW
EPDQTFSLNRDLTGELEYATKISRFSNVYHLE HISKNFGADTTKVFYIGLRGEWTELRRHEVT CNYEASANPADHRVHQVTPQTHFIS 554 1904 A 4200 1 961 GIPCTEMGNFDNANVTGEIEFAIHYCFKTHSI EICIKACKNLAYGEEKKKKCNPYVKTYLLPI RSSQGKRKTGVQRNTVDPIFQETLKYQVAP QLVTRQLQVSVWHLGTLARRVFLGEVIIPLA WDFEDSTTQSFRWHPLRAKADK YEDSVPQS NGELTVRAKLVLPSRTRKLQEAQEGTDQPSI HGQLCLVVLGAKNLPVRPDGTLNSFVKGCL LPDQQKLRLKSPVLRKQACPQWKHSFVFSG TPAQLRQSSLELTVWDQALFGMNDRLLGGT RLGSKGDTAVGGDACSQSKLQWQKVLSSPN LWTDMTLVLH 555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERR WQQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKI QLASQPDLRDGTNYRDSEVPRESSHEDSLLE		ľ			1			ERAD\DRSKFVESDADEELLFNIPFTG\HVKLK
HISKNFGADTTKVFYIGLRGEWTELRRHEVT CNYEASANPADHRVHQVTPQTHFIS 554 1904 A 4200 1 961 GIPCTEMGNFDNANVTGEIEFAIHYCFKTHSI EICIKACKNLAYGEEKKKKCNPYVKTYLLPI RSSQGKRKTGVQRNTVDPIFQETLKYQVAP QLVTRQLQVSVWHLGTLARRVFLGEVIIPLA WDFEDSTTQSFRWHPLRAKADKYEDSVPQS NGELTVRAKLVLPSRTRKLQEAQEGTDQPSI HGQLCLVVLGAKNLPVRPDGTLNSFVKGCL LPDQQKLRLKSPVLRKQACPQWKHSFVFSG TPAQLRQSSLELTVWDQALFGMNDRLLGGT RLGSKGDTAVGGDACSQSKLQWQKVLSSPN LWTDMTLVLH 555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERR WQQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKI QLASQPDLRDGTNYRDSEVPRESSHEDSLLE	ı							GIIIMGEDDDSHPSEMRLYKNIPQMSFDDTER
CNYEASANPADHRVHQVTPQTHFIS							.]	EPDQTFSLNRDLTGELEYATKISRFSNVYHLSI
1904 A 4200 I 961 GIPCTEMGNFDNANVTGEIEFAIHYCFKTHSJEICIKACKNLAYGEEKKKKCNPYVKTYLLPIRSSQGKRKTGVQRNTVDPIFQETLKYQVAPQLVTRQLQVSVWHLGTLARRVFLGEVIIPLAWDFEDSTTQSFRWHPLRAKADKYEDSVPQSNGELTVRAKLVLPSRTRKLQEAQEGTDQPSIHGQLCLVVLGAKNLPVRPDGTLNSFVKGCLLPDQQKLRLKSPVLRKQACPQWKHSFVFSGRLGSKGDTAVGGDACSQSKLQWQKVLSSPLWTDMTLVLH 555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDHNHHENERRWQQERLHREEAYYQFINELNDEDYRLMRDHNLLGTPGEITSEELQQRLDGVKIQLASQPDLRDGTNYRDSEVPRESSHEDSLLE	1	ł	i					HISKNFGADTTKVFYIGLRGEWTELRRHEVTI
EICIKACKNLAYGEEKKKKCNPYVKTYLLPI RSSQGKRKTGVQRNTVDPTFQETLKYQVAP QLVTRQLQVSVWHLGTLARRVFLGEVIIPLA WDFEDSTTQSFRWHPLRAKADKYEDSVPQS NGELTVRAKLVLPSRTRKLQEAQEGTDQPSI HGQLCLVVLGAKNLPVRPDGTLNSFVKGCL LPDQQKLRLKSPVLRKQACPQWKHSFVFSG TPAQLRQSSLELTVWDQALFGMNDRLLGGT RLGSKGDTAVGGDACSQSKLQWQKVLSSR LWTDMTLVLH 555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKI QLASQPDLRDGTNYRDSEVPRESSHEDSLLE	\vdash	554	1004		4200		061	
RSSQGKRKTGVQRNTVDPIFQETLKYQVAP QLVTRQLQVSVWHLGTLARRVFLGEVIIPLA WDFEDSTTQSFRWHPLRAKADK YEDSVPQS NGELTVRAKLVLPSRTRKLQEAQEGTDQPSI HGQLCLVVLGAKNLPVRPDGTLNSFVKGCL LPDQQKLRLKSPVLRKQACPQWKHSFVFSG TPAQLRQSSLELTVWDQALFGMNDRLLGGT RLGSKGDTAVGGDACSQSKLQWQKVLSSR LWTDMTLVLH 555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERR WQQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKI QLASQPDLRDGTNYRDSEVPRESSHEDSLLE	13	JJ4	1304	A	4200	1	וסע	
QLVTRQLQVSVWHLGTLARRVFLGEVIIPLA WDFEDSTTQSFRWHPLRAKADKYEDSVPQS NGELTVRAKLVLPSRTRKLQEAQEGTDQPSI HGQLCLVVLGAKNLPVRPDGTLNSFVKGCL LPDQQKLRLKSPVLRKQACPQWKHSFVFSG TPAQLRQSSLELTVWDQALFGMNDRLLGGT RLGSKGDTAVGGDACSQSKLQWQKVLSSP LWTDMTLVLH 555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKI QLASQPDLRDGTNYRDSEVPRESSHEDSLLE		- 1	ł					
WDFEDSTTQSFRWHPLRAKADK YEDSVPQS NGELTVRAKLVLPSRTRKLQEAQEGTDQPSI HGQLCLVVLGAKNLPVRPDGTLNSFVKGCL LPDQQKLRLKSPVLRKQACPQWKHSFVFSG TPAQLRQSSLELTVWDQALFGMNDRLLGGT RLGSKGDTAVGGDACSQSKLQWQKVLSSP\ LWTDMTLVLH 555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKI QLASQPDLRDGTNYRDSEVPRESSHEDSLLE		. !				•		
NGELTVRAKLVLPSRTRKLQEAQEGTDQPSI HGQLCLVVLGAKNLPVRPDGTLNSFVKGCL LPDQQKLRLKSPVLRKQACPQWKHSFVFSG TPAQLRQSSLELTVWDQALFGMNDRLLGGT RLGSKGDTAVGGDACSQSKLQWQKVLSSPN LWTDMTLVLH 555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERR WQQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKI QLASQPDLRDGTNYRDSEVPRESSHEDSLLE	1					ĺ		
HGQLCLVVLGAKNLPVRPDGTLNSFVKGCL LPDQQKLRLKSPVLRKQACPQWKHSFVFSG TPAQLRQSSLELTVWDQALFGMNDRLLGGT RLGSKGDTAVGGDACSQSKLQWQKVLSSP\ LWTDMTLVLH 555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERR WQQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKI QLASQPDLRDGTNYRDSEVPRESSHEDSLLE								
LPDQQKLRLKSPVLRKQACPQWKHSFVFSG TPAQLRQSSLELTVWDQALFGMNDRLLGGT RLGSKGDTAVGGDACSQSKLQWQKVLSSPN LWTDMTLVLH 555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERR WQQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKI QLASQPDLRDGTNYRDSEVPRESSHEDSLLE		ŀ					ŀ	
TPAQLRQSSLELTVWDQALFGMNDRLLGGT RLGSKGDTAVGGDACSQSKLQWQKVLSSPN LWTDMTLVLH 555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKI QLASQPDLRDGTNYRDSEVPRESSHEDSLLE		1	ľ	' i		ł		
RLGSKGDTAVGGDACSQSKLQWQKVLSSPYLWTDMTLVLH 555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKI QLASQPDLRDGTNYRDSEVPRESSHEDSLLE		l		J				
LWTDMTLVLH	1	{		[
555 1905 A 4211 331 2419 KENKKARNLRIMNQSRSRSDGGSEETLPQDH NHHENERR WQQERLHREEA YYQFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKI QLASQPDLRDGTNYRDSEVPRESSHEDSLLE	-	į		- 1				
NHHENERR WQQERLHREEA YYQFINELNDE DYRLMRDHNLLGTPGEITSEELQ QRLDGVKI QLASQPDLRDGTNYRDSEVPRESSHEDSLLE	5	555	1905	A	4211	331	2419	
DYRLMRDHNLLGTPGEITSEELQQRLDGVKI QLASQPDLRDGTNYRDSEVPRESSHEDSLLE		1	{	[NHHENERRWQQERLHREEAYYQFINELNDE
QLASQPDLRDGTNYRDSEVPRESSHEDSLLE		İ		1				DYRLMRDHNLLGTPGEITSEELQQRLDGVKE
		1	ļ	- 1	1			QLASQPDLRDGTNYRDSEVPRESSHEDSLLE
The state of the s		1	1	- (ĺ			WLNTFRRTGNATRSGQNGNQTWRAVSRTNP
		i			}	}	j	NNGEFRFSLEIHVNHENRGFEIHGEDYTDIPLS
					1			DSNRDHTANRQQRST\SPVARRTRSQTSVNFN
GSSSNIPRTRLASRGQNPAEGSFSTLGRLRNG	L		l					GSSSNIPRTRLASRGQNPAEGSFSTLGRLRNGI

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion GGAAGIPRANASRTNFSSHTNQSGGSELRQRE GQRFGAAHVWENGARSNVTVRNTNQRLEPI RLRSTSNSRSRSPIQRQSGTVYHNSQRESRPV
						QQTTRRSVRRRGRTRVFLEQDRERERRGTAY TPFSNSRLVSRITVEEGEESSRSSTAVRRHPTIT LDLQVRKIRPGENRDRDSIANRTRSRVGLAE NTVTIESNSGGFRRTISRLERSGIRTYVSTITVP LRRISENELVEPSSVALRSILRQIMTGFGELSSL MEADSESELQRNGQHLPDMHSELSNLGTDN NRSQHREGSSQDRQAQGDSTEMHGENETTQP HTRNSDSRGGRQLRNPNNLVETGTLPILRLAH FFLLNESDDDDRIRGLTKEQIDNLSTRHYEHN SIDSELGKICSVCISDYVTGNKLRQLPCMHEF HIHCIDRWLSENCTCPICRQPVLGSNIANNG
556	1906	Α	4212	3	462	LQRQRQHPAAAPAVPVRCFTFCFTDIVIMPKR KSPENTEGKDGSKVTKQEPTRRSARLSAKPA PPKPEPKPRKTSAKKEPGAKISRGAKGKKEEK QEAGKEGTAPSENGETKAEEIHISRSTVNVST SRGTPPSTLSVKGQIETVRVKGTEN
557	1907	A	4213	774	507	ARRFSCLTLQTSWGHRH\GPPRP\ANFVFLVET GFLHIGQAGHKLPTSGDPPASASQSARITGMS HRTWFLASFLIDSCKNFIVYKIMYTL
558	1908	A	4225	3	1253	TYRHAEREHPETSSATKVSYDYRHKRPKLLD GDQDFSDGRTQKYCKEEDRKYSFQKGPLNRE LDCFNTGRGRETQDGQVKEPFKPSKKDSIAC TYSNKNDVDLRSSNDKWKEKKKEGDCRKE SNSSSNQLDKSQKLPDVKPSPINLRKKSLTVK VDVKKTVDTFRVASSYSTERQMSHDLVAVG RKSENFHPVFEHLDSTQNTENKPTGEFAQEIIT IIHQVKANYFPSPGITLHERFS\KMADIHKADV NEIPLNSDPEIHRRIDMSLAELQSKQAVIYESE QTLIKIIDPNDLRHDIERRRKERLQNEDEHIFHI ASAAERDDQNSSFSKNYTTQRKDIITHKPFEV EGNHRNTRVRPFKSNFRGGRCQPNYKSGLVQ KSLYIQAKYQRLRFTGPRGFITHKFRERLMRK KKVP
559	1909	A	4235	1	323	KFSIPFFLRWSFTLV\PRLEGNDMISVHCNLGL LGLSHSPASASQVGGITGTQHHTGLIFGFLIET EFHHVGQAGLELLTSGDPPALAFQSAGITGVS HHAWLQVLNS
560	1910	A	4246		1569	TLSLLERVLMKDIVTPVPQEEVKTVIRKCLEQ AALVNYSRLSEYAKIEGKKREMYELPVFCLA SQVMDLTIQNQKDAENVGRLITPAKKLEDTIR LAELVIEVLQQNEEHHAEAFAWWSDLMVEH AETFLSLFAVDMDAALEVQPPDTWDSFPLFQ LLWDFLRTGLLICGNGK\FHKHLQDLFAPLVV R/YMWDLDGSSPIAQSIHRGLLSRESWEPVNN GSGTSEDLFWKLDALQTFIRDLHWPEEEFGK HLEQRLKLMASDMIESCVKRTR\IAFEVKLQK TSSIQQIFRVPQFNMAPCFNVMGLMAKGSIQP KL\CSMEMGQEFAKMWHQYHSKIDELIEETV KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS SFLSFTVKAASKYVDVPKPGMDVADAYVTF VRHSQDVLRDKVNEEMYIERLFDQWYNSSM NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY RDFRLQGVLDSTLNSKTYETIRNRLTVEEATA SVSEGGGLQGISMKDSDEEDEEDD
561	1911	A	4257	1300	654	SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI FPDLFKRAAERLQYVFGYKLVELEPKSNTYIL

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-		nou	1	nucleotide		
1	peptide	l	in	1	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	ĺ	ĺ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
			1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	į		!	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			1	peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
			i			INTLEPVEEDAEMRGDQGTPTTGLLMIVLGLI
			İ		ľ	FMKGNTIKETEAWDFLLAL\GVYPTKKHLIFG
			i		i	DPKKLITEDFVRORYLEYRRIPHTDPVDYEFO
ł		1	i			WGPRTNLETSKMKVLKFVAKVHNQDPKDW
			1			
662	1010	 	1000	 	1400	PAQYCEALADEENRARPQPSGPAPSS
562	1912	Α	4260	1	1498	MVTWLYRFLPTSNMAAKLRSLLPPDLRLQF
1			1	į –	:	WLHARLQKCFLSRGCGSYCAGAKASPLPGK
Į			i			MAMGLMCGRRELLRLLQSGRRVHSVAGPSQ
1		l	į			WLGKPLTTRLLFPAAPCCCRPHYLFLAASGPR
1			ļ			SLSTSAISFAEVQVQAPPVVAATPSPTAVPEV
1			1			ASGETADVVQTAAEQSFAELGLGSYTPVGLI
			!			QNLLEFMHVDLGLPWWGAIAACTVFARCLIF
1		1	i	l		PLIVTGQREAARIHNHLPEIQKFSSRIREAKLA
1		1		· ·		GDHIEYYKASSEMALYQKKHGIKLYKPLILPV
			Į.	[TQAPIFISFFIALREMANLPVPSLQTGGLWWF
			!			1 -
		1	1	[ĺ	QDLTVSDPIYILPLAVTATMWAVLELGAETG
1						VQSSDLQWMRNVIRMMPLITLPITMHFPTAV
1.		İ				FMYWLSSNLFSLVQVSCLRIPAVRTVLKIPQR
ļ						VVHDLDKLPPREGFLESFKKGWKNAEMTRQ
1			}	1	1	LREREQRMRNQLELAARGPLRQTFTHNPLLQ
[<u> </u>			ĺ		PGKDNPPNIPSS\SSSSSKPKSKYPWHDTLG
563	1913	Α	4265	623	116	MGGLAPTQTLEPT\REYQNTQLSVSYLLPEQN
1						THGTRRTLSSGPSNNLPLPLSSSATMPSMOCK
1			1	[HRSPNGGLFRQSPVK/TPPIPMSFQPVPGGV\L
1			1	[ļ	PRGSGNPPHGTSILTAPPALLPHPPTHPTQQSF
1 :		1	i	İ		LIQENNNTNHTHSHTHTYTETLSFFLYICVNN
			ļ			DRMEWGKSVF
564	1914	A	4270	3	368	
304	1914	^	4270)	300	ILKRKLSSLNSEVSTIQNTRMLAFKATAQLFIL
			1			GCTWCLGLLQVGPAAQVMAYLFTIINSLQGF
1	{		Ī	!	ľ	FIFLVYCLLS\QQVQKQYQKWFREIVKSKSES
<u>L</u>						ETYTLSSKMGPDSKPSEGDVFPRTSE
565	1915	A	4288	83	406	RNSRPL WCSPPASQPRQAPVSQSCCCPLPSSSS
∤ . i			1			PPSALLAPTKPRALGTLRLYECSPELCTTMLP
						PAWLLMLCQAPRPQDPDPRLTQPEKSLQEAP
						GOTGASRTPRT
566	1916	A	4298	1041	229	LNSSQKLACLIGVEGGHSLDSSLSVLRSFYVL
				· · · · ·	/	GVRYLTLTFTCSTPWAESSTKFRHHMYTNVS
		}				GLTSFGEKVVEELNRLGMMIDLSYASDTLIRR
]		1	
					.	VLEVSQAPVIFSHSAARAVCDNLLNVPDDILQ
1					1	LLKKNGGIVMVTLSMGVLQCNLLANVSTVA
1				•		DHFDHIRAVIGSEFIGIGGNYDGTGRFPQGL\E
}		}				DVSTYPVLIEELLSRSWSEEELQGVLRGNLLR
]		l				VFRQVEKVREESRAQSPVEAEFPYGQLSTSCH
		1				FHLGASEWTPRLLIWR
567	1917	Α	4299	1	1106	GATPLGSVGGRTGKMDAATLTYDTLRFAEFE
		1		İ		DFPETSEPVWILGRKYSIFTEKDEILSDVASRL
]			•	į		WFTYRKNFPAIGGTGPTSDTGWGCMLRCGQ
					,	MIFAQALVCRHLGRDWRWTQRKRQPDSYFS
						VLNAFIDRKDSYYSIHQIAQMGVGEGKSIGQ
[
[.]						WYGPNTVAQVLKKLAVFDTWSSLAVHIAMD
					i	NTVVMEEIRRLCRTSVPCAGATAFPADSDRH
						CNGFPAGAEVTNRPSPWRPLVLLIPLRLGLTD
						INEAYVETLKHCFMMPQSLGVIGGKPNSAHY
1						FIGYVGEELIYLDPHTTQPAVEPTDGCFIPDES
						FHCQHPPCRMSIAELDPSIAVVRGGHLSTQAF
					!	GAECCLGMTRKTFGFLRFFFSMLG
568	1918	Α	4300	2012	1843	SRKFLTITPIVLYFLTSFYTKYDQIHFVLNTVS
555	1710	1.	1,500	2012	1070	
569	1010	<u> </u>	1200	106	521	LMSVLIPKLPQLHGVRIFGINKY
	1919	A	4302	186	531	WTFCLFL/WWVPESARWLLTQGHVKEAHRY

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LLHCARLNGRPVCEDSFSQEVRVNVCVSMHI CVWWGVGCVKCLPPRAHHIWQEKPLGPHRT
570	1920	A	4308	3	869	VTESKLEAEGKTKEKAREKERKKKS RSGQGKVYGLIGRRRFQQMDVLEGLNLLITIS GKRNKLRVYYLSWLRNKILHNDPEVEKKQG WTTVGDMEGCGHYRVVKYERIKFLVIALKSS VEVYAWAPKPYHKFMAFKSFADLPHRPLLV DLTVEEGQRLKVIYGSSAGFHAVDVDSGNSY DIYIPVHIQSQITPHAIIFLPNTDGMEMLLCYE DEGVYVNTYGRIIKDVVLQWGEMPTSVAYIC SNQIMGWGEKAIEIRSVETGHLDGVFMHKRA QRLKFLCERNDKVFFASVRSGGSSQVYFMTL NRNCIMNW
571	1921	A	4309	9	524	ASREMDVTKVCGEMRYQLNKTNMEKDEAE KEHREFRAKTNRDLEIKDQEIEKLRIELDESK QHLEQEQQKAALAREECLRLTELLGESEHQL HLTRQEKDSIQQSFSKEAKAQALQAQQREQE LTQKIQQMEAQHDKTENEQYLLLTSQNTFLT KLKEECCTLAKKLEQISQ
572	1922	A	4318	1	1119	GATPLGSVGGRTGKMDAATLTYDTLRFAEFE DFPETSEPVWILGRKYSIFTEKDEILSDVASRL WFTYRKNFPAIGGTGPTSDTGWGCMLRCGQ MIFAQALVCRHLGRDWRWTQRKRQPDSYFS VLNAFIDRKDSYYSIHQIAQMGVGEGKSIGQ WYGPNTVAQVLKKLAVFDTWSSLAVHIAMD NTVVMEEIRRLCRTSVPCAGATAFPADSDRH CNGFPAGAEVTNRPSPWRPLVLLIPLRLGL\T DINEAYVLETL\KHCFHGWPQFPG/VVHREGK PNSAHYFIGYVGEELIYLDPHTTQPAVEPTDG CFIPDESFHCQHPPCRMSIAELDPSIAVVRGGH LSTQAFGAECCLGMTRKTFGFLRFFFSMLG
573	1923	A .	4333	363	1066	GGVPVGLASKPFQILYGHTNEVLSVGISTELD MAVSGSRDGTVIIHTIQKGQYMRTLRPPCESS LFLTIPNLAISWEGHIVVYSSTEEKTTLKÆRM HYICFSINGKYLGSQILKEQVSDICIIGEHIVTG SIQGFLSIRDLHSLNLSINPLAMRLPIHCVCVT KEYSHILVGLEDGKLIVVGVGKPAEVKPSISN FISHAVGDYFGSPSFQLIEKSPLGINKLKAKFD FSKGSK
574	1924	A	4346	359	1234	MDTLEEVTWANGSTALPPPLAPNISVPHRCLL LLYEDIGTSRVRYWDLLLLIPNVLFLIFLLWK LPSARAKIRITSSPIFITFYILVFVVALVGIARA VVSMTVSTSNAATVADKILWEITRFFLLAIEL SVIILGLAFGHLESKSIKRVLAITTVLSLAYSV TQGTLEILYPDAHLSAEDFNIYGHGGRQFWL VSSCFFFLVYSLVVILPKTPLKERISLPSRRSFY VYAGILALLNLLQGLGSVLLCFDIEGLCCVD ATTFLYFSFFAPLIYVAFLRGFFGSEPKILF
575	1925	A	4360	2038	1512	GCWWRHPWLASQRDCLDCRIQLAEKFVKAV SKPSRPDMNPIRVKEVYRLEEMEKIFVRLEM KIIKGSSGTPKLSYTGRDDRHFVPMGLYIVRT VNEPWTMGFSKSFKKKFFYNKKTKDSTFDLP ADSIAPFHICYYGRLFWEWGDGIRVHDSQKP QDQDKLSKEDVLSFIQMHRA
576	1926	A	4365	69	500	QVEGRQGREVKRTA WRISPVWRPARCRRRST PQP/PE/PGAQQQERHRQGEAPMQALDPRAEP GPQAQSHAACQPEPEPPRVLLDPTAARGGVQ GRP/GLSRHPGLAPHPQTHTPWPQSGRLPCAS EPLPLGGIRPTPGLEPKGRDLM

SEQ ID	SEQ ID	Met	SEQ	Predicted	Dradioted and	Amino acid comence (AmAlonia C. Curtai
NO: of	NO: of	hod	ID NO:	beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid.
nucl-	peptide	1100	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	ļ			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	1		1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	!		}	peptide	1	/=possible nucleotide deletion, \=possible
		<u></u>		sequence	<u></u>	nucleotide insertion
577	1927	A	4366	785	502	SAPPKKKNGVLFLSPRLKSSGAIWVHSTPTLW
1	[1 :		İ	ASSNSRASTPKVAGITGARPHARIIFVFLIEMG
570	1000	<u> </u>	1			FHNVGQAGL/DTLTLVICPPQPPKLLGLQM
578	1928	A	4367	1	221	FFFFLKKSRCVTQAGVQG\PISLHPPPPGFKRF
1			ļ i		1	SRLSLLSSWDYRHP/HAANFCIFSRDG\VSPYW
579	1929	A	4383	1	224	SGWSRTPDLR FETESHSVTOAGMOWUNII GSLOPMPRGLER
1 3/3	1,729	^	4303	•	444	FETESHSVTQAGMQWHNLGSLQPMP/PGLKR FSCLRLQSSWDHRHAPPHLAHFCIFSRDGVSP
1					1	CWPGWSSTPDLK
580	1930	A	4397	410	94	SRLKPYSTNVTAKKLPATNIPNLDCFTAKLYO
			.57,		ļ ⁻ .	VFKKGNHILHELFQNKEEGAFPNS/FYEASFT
1	1		1			LRPKSDRDIAKEESYSTISLLSTDTKILMSKYK
L	l	·	1			QLKSSDL .
581	1931	Α	4414	670	3 .	VLVHRQCGGILRLRRKEAVSVLDSADIEVTDS
1	Į.		1			RLPHATIVDHRPQHRWLETCNAPPQLIQGKA
j i	ļ		J i]	RSAPKPSQASGHFSVELVRGYAGFGLTLGGG
1	•	1		ľ	1	RDVAGDTPLAVRGLLKDGP\AQRCGRLEVGD
4	1	l				LVLHINGESTQGLT\HAQAVERIRAGGPQLHL
i	1 ,	1			ł	VIRRPLETHPGKPRGVGEPRKGVVPSWPDRSP
4	· 1	l)			DPGGPEVTGSRSSSTSLVQHPPSRTTLKKTRG
582	1932	A	4424	194	449	SPE VLYIRKKKRLEKLRHQLMPMYNFDPTEEQDE
302	1752	^	***	1,77	777	LEQELLEHGRDAASVQAATSVQAMQGKTTL
1	†	1		·		PS\QGPLQRPSRLVFT\DVANAIHV
583	1933	A	4435	1	166	APGPPVPPPGSPPEQMPGPCPASMPP/DPPPGS
	l				- 1-	PPEQMPGPCPVSAPP/GPPPGSPPEQMPGPCPV
L \	<u></u>		<u> </u>			SAPPALLQDTSV
584	1934	A	4439	1	628	SATPQQPSAPQHQGTLNQPPVPGMDESMSYQ
	ļ ,	ļ	j . j			APPQQLPSAQPPQPSNPPHGAHTLNSGPQPGT
1	۱ ۱	Į				APATQHSQAGPATGQAYGPHTYTEPAKPKK
 			 			GQQLWNRMKPAPGT\EVSSSTSRSDPLLLPPR
l l	١ ،	ł		1		ALAPTQRASTVVLAPSPT/SEKVQNHSGSSAR
	۱ ۱					GNLSGKPDDWP/LGHERVCGALLHRL*VGGG QGPHGKAAQGGAAGAAAGRLGLYH
585	1935	Α	4463	10	144	HKPVTNSRDTQEVPLEKAKQVLKIIATFKHTT
	1	١	.,,,,			SIFDDFAHYEKRO
586	1936	A	4464	1309	103	LNAESYVSFTTKLDIPTAAKYEYGVPLOTSDS
	' i	١ ,	"		-	FLRFPSSLTSSLCTDNNPAAFLVNQAVKCTRK
	¹ i	١ ,	į 1			INLEQCEEIEALSMAFYSSPEILRVPDSRKKVPI
	¹	١ ,				TVQSIVIQSLNKTLTRREDTDVLQPTLVNAGH
	¹ . i	١ ,				FSLCVNVVLEVKYSLTYTDAGEVTKADLSFV
]	' i	۱ ۱				LGTVSSVVVPLQQKFEIHFLQENTQPVPLSGN
	1	۱ ۱	, 1		!	PGYVVGLPLAAGFQPHKGSGIIQTTNRYGQLT
1		۱. ا	Į l	l		ILHSTTEQDCLALEGVRTPVLFGYTMQSGCK
1	1	! 1	ı l			LRLTGALPCQLVAQKVKSLLWGQGFPDYVA PFGNSOGP/ADMLDWVPIHFITOSFNRKDSCO
1		, i	ļ		1	LPGALVIEVKWTKYGSLLNPQAKIVNVTANLI
1	1	۱ ۱	ı l			SSSFPEANSGNERTILISTAVTFVDVSAPAEAG
_ 1	· }	۱ ۱	ı l			FRAPPAINARLPFNFFFPFV
587	1937	A	4471	614	387	LLGRASAC/LQLQSSW/D/HRPMLPYLANFVF
}	1	۱ ۱	ı l		·	CKDR/SFTWLPRLVLNSWLQVILLPWPPTGCD
		<u> </u>	<u> </u>			NKHEPPCPATKRRHSGSI
588	1938	A	4480	1720	1458	HDLGSLQPPPPGFKRFSCLSLPSSWDYRLMPP
1		۱ ۱	ļ †			CPANFCIIII/DFLVETGFHHVGQASHELLTSGD
500	1022	اا	1	000		PPTSASQSAGITGMSYHTWFGES
589	1939	A	4487	922	332	APVTTSPRVGQPW/RTALALRSLYRARPSLRC
1	' I	' I	ı l		}	PPVELPWAPRRGHRLSPADDELYQRTRISLLQ
	' I	' I	(I			REAAQAMYIDSYNSRGFMINGNRVLGPCALL PHSVVOWNVGSHODITEDSESI FWLLEPRIEI
		'i	L			PHSVVQWNVGSHQDITEDSFSLFWLLEPRIEI

NO: of No: of note not	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Decidide Sequence			l .				
	1)			
Sequence					1		
uence	1					to last amino	
amino acid residue of peptide residue of peptide residue of peptide sequence T=Thronine, V=Juline, W=Typtophan, Y=Typt				1			
Presidue of peptide	denies			111			
peptide	1		ľ	ł			
	1			1		sequence	
1940 A 4492 1 472		1	ĺ				
DTPMACATINELCHEGRYTGAALIPPEGGTSL TSLGQAQQ TSLGAQAQAQ TSLGA	<u> </u>			 -	sequence	ļ	
1940 A 4492 1 472 FFFEEESRSYAQAGVQWEDLGSLQAPPPGFT FFFSCLSLPSSWOTRPPLRANFEVEL VETICEP FFFSCLSLPSSWOTRPPLRANFEVEL VETICEP FFSSCLSLPSSWOTRPPLRANFEVEL VETICEP FFSSCLSLPSSWOTRPPLRANFEVEL VETICEP FFSSCLSLPSSWOTRPPLRANFEVEL VETICEP FFSSCLDLLT/SIGDPTSASGQAGTGVSHR APPRIGEPRIKGOAAVVWPSTSLGDHRVTS VPHQGGLPGPIRVAPSSAGQREASQCPPGR APPRIGEPRIKGOAAVVWPSTSLGDHRVTS VPHQGGLPGPIRVAPSSAGQREASQCPPGR IA44 III6 IAARPTLAKTWNQLKRPTMJDSIKKTRYTYT MEY ADTERNEMSFAGT WYBLEAILLSKLM KLDNWYBLTPIQGAVCTATAGGMKRILFAL FEWDSSCPHPSSGV FFSGCSCPHPSSGV FFFGCSCPTPAST FFFGCSCPTPAST FFFGCSCPTPAST FFFGCSCPTPAST FFFGCSCPTPAST FFFGCSCPTPAST FFFGCSCPTPAST FFFGCSCPTPAST FFFGCSCPTPAST FFFGCSCSPQAGVQRCWMFDLFFILOR FFFGCSCSPQAGVQRTTHABHTQLFTALFQTT AVQVCRQADIRAFTYPTIV VEDAGGKRSSJ FFFGASSCSVPQAGVQRTHABHTQLFTALFQT FFFGASSCSVPQAGVQRTHABHTQLFTALFQT FFFGASSCSVPQAGVQRTDLGWPDV FFFGASSCSVPQAGVQRTHABHTQLFTALFQT FFFGASSCSVPQAGVQRTHABHTQLFTALFQT FFFGASSCSVPQAGVGTTHABHTQLFTALFQT FFFGASSCSVPQAGVGTTHABHTQLFTALFQT FFFGASSCSVPQAGVGTTHABHTQLFTALFQT FFFGASSCSVPQAGVGTTHABHTQLFTALFQT FFFGASSCSVPQAGVGTTHABHTQLFTALFQT FFFGASSCSVPQAGVGTTHABHTQLFTALFQT FFFGASSCSVPQAGVGTTHABHTQLFTALFQT FFFGASSCSVPQAGVGTTHABHTQLFTALFQT FFFGASSCSVPQAGVGTTHABHTQLFTALFQT FFFGASSCSVPQAGVGTTHABHTQLFTALFQT FFFGASSCSVPQAGVGTTHABHTQLFTALFQT FFFGASSCSVPQAGVGTTHABHTQLFTALFQT FFFGASSCSVPQAGVGT FFFGASSCSVPQAGVGT FFFGASSCSVPCAGVGT FFFGASSCSVPQ				1			DTDMACATEMEN CHECONTECA AL INDICATEN
1940 A 4492 1 472			l	ļ			
PFSCLSLPSSWDYRRPPLRPAMFFVFLVETGFP RPSRDGIDLITIGOPPTASOAGTOVSIR ARPKRIGEPTRSCORDPTASOAGTOVSIR ARPKRIGEPTRSCORDPTASOAGTOVSIR ARPKRIGEPTRSCORDPTASOAGTOPPTASOAG	500	1040	_	4400	1	470	
	1 390	1940	A	4492	1	4/2	
SPENDED			,	l)	1	
S91	ĺ				Í		
1941	,						
MEYYADTERNEIMSPAGTWVELEAIILSKI.M LKDNWVEDTIPQAVPCATAEGMKRILFAL	501	1041	ļ_ ,	1105			
LKDNWVEDTIPQGAVPCTATAEGMKRLLFAL	391	1941	A	4495	1444	1116	
	i l						
1942 A							
VCGGLSLANAWGILSVGAKQKKWKPLEFI							
LCTLAATHM.NVAVPIATYSVVQLRGRPDF	592	1942	Α	4496	2	919	RTRPLFSGRPTRPVCTMSDERRLPGSAVGWL
BWNEGLCKYFYSTFYTLTLATCTSYTSLSYHR MWMVCWPVNYRLSNAKKQAGHTVMGIWM GSFILSALPAVQWHDTSERFYTHGCRFIVAE GLGFGVCFLLLVGGSVAMGVICTAILFQTI]			ľ	ĺ		VCGGLSLLANAWGILSVGAKQKKWKPLEFL
							LCTLAATHMLNVAVPIATYSVVQLRRQRPDF
GSFILSALPAVGWHDTSERFYTHGCRFIVAEL GLGFGVCFLLLVGGSVAMGVICTALAFQTL AVQVGRQADHRAFTYTTVDCMGFVL AVQVGRQADHRAFTYTTVTDCMGFVL GPFSLADTHLSDLPYTTVFVDCLMGFVL GPFSLADTHLSDLPYTHGDRDSGGACVM GPFSLADTHLSDLPYTHGDRDSGGACVM GPFSLADTHLSDLPYTHGDRDSGGACVM GPFSLADTHLSDLPYTHGDRDSGGACVM GPFSLADTHLSDLPYTHGDRDSGGACVM GPFSLADTHLSDLPYTHGDRDSGGACVM GPFSLADTHLSDLPYTHGDRDSGGACVM GPFSLADTHLSDLPYTHGDRDSGGACVM GPFSLADTHLSDLPYTHGDRDSGGACVM GPFSLADTHLSDLFTHAFT GPFSLADTHLSDLFTHAFT GPFSLADTHLSDLFTHAFT GPFSLADTHLSDLFTHAFT GPFSLATT GPF							EWNEGLCKVFVSTFYTLTLATCFSVTSLSYHR
GIGFGVCFLILVGGSVAMGVICTAIALFOTI							MWMVCWPVNYRLSNAKKQAGHTVMGIWM
A	1						GSFILSALPAVGWHDTSERFYTHGCRFIVAEI
DOSEPAKTSLQTTGLVTTIVFIYDCLMGPPVL GFFSLADTHLSDLPTWGDRDSGACVM	i i						GLGFGVCFLLLVGGSVAMGVICTAIALFOTL
DOSEPAKTSLQTTGLVTTIVFIYDCLMGPPVL GFFSLADTHLSDLPTWGDRDSGACVM							
GFFSLADTHLSDLPYTWGDRDSGGACVM	[[
1943							
HFPASASQVAGTTHARHHTQLIF\AFILVENGL C	593	1943	A	4506	2	193	
1944							
ESTEIPPYVMKCPSNGLCSRLPADCIDCTTNPS CTYGKPVTIPCAVKPSVTCVDQDFKSQKNFII NMTCRFCWQLPFTDYECTNSTSCMTVSCPRQ RYPANCTVRUDIVHCLGNRTFPKMLYCNWT GGYKWVYGLWLIRHIPRWGLGADRFYLGP VAGTASGKLFSFGGLGIWTLIDVLLIGVGYVG PADOSLYI 595 1945 A 4512 533 264 FFFKMESYSVARLECSGAISAPCNLHLLGSNN SPASASRVIAGNIGARHHTQQIPVLLVQMRVH YVGQDGLDLLNLMHPPRSPKVLGLQA 596 1946 A 4513 3 1674 HASDHLYPNFLVNELIIKQKQRFEEKRFKLD HSVSSTNGHRWQIPQDWLGTDQDNLDLANV NLMLELLVQKKKQLEAESHAAQLQILMEFLK VARRNKREQLEQIQKELSVLEEDIKRVEEMS GLYSPVSEDSTVPQFEAPSPSHSSIIDSTEYSQP PGFSGSSQTKKQPWYNSTLASRRKRLTAHFE DLEQCYFSTRMSRISDDSRTASQLDEFQECLS KFTRYNSVRPLATLSYASDLYNGSQYKSLV FFEFDRDCDYFAIAGVTKKIKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHLAFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKPYCLRSFKGHIN EKNFVGLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANSQGTRKVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHITVLIFCILGQMGFHHVGE QAGLELLTSWHITVLIFCILGQMGFHHVGE QAGLELLTSWHITVLIFCILGQMGFHHVGE QAGLELLTSWHITVLIFCILGQMGFHHVGE QAGLELTSWAGGTISVRHITVLIFCILGQMGFHHVGE QAGLELTSWAGGTISVRHITVLIFCILGQMGFHHVGE QAGLELTSWAGGTISVRHITVLIFCILGQMGFHHVGE QAGLELTSWAGGTISVRHITVLIFCILGQMGFHHVGE QAGLELTSWAGGTISVRHITVLIFCILGQMGFHHVGE QAGLELTSWAGGTISVRHITVLIFCILGQMGFHHVGE QAGLELTSWAGGTISVRHITVLIFCILGQMGFHHVGE QAGLELTSWAGGTISVRHITVLIFCILGQMGFHHVGE QAGLELTSWAGGTISVRHITVLIFCILGQMGFHAGE SPR	l i						
ESTEIPPYVMKCPSNGLCSRLPADCIDCTTINFS CTYGKPVTFDCAVKPSVTCVDQDFKSQKNFII NMTCRFCWQLPETDYECTNSTSCMTVSCPRQ RYPANCTVRIDHYHCLGNRTFFKMLYCNWT GGYKWVYGLWLLRHHPRWGLGADRF\YLGP VAGTASGKLFSFGGLGIWTLIDVLLIGVGYVG PADGSLYI 595 1945 A 4512 533 264 FFFKMESYSVARLECSGAISAPCNLHLLGSNN SPASASRV/AGNIGARHHTQQIFVLLVQMRVH YVGQDGLDLL/NLMIHPPRSPKVLGLQA 596 1946 A 4513 3 1674 HASDHLYPNFLVNFILIKQKQRFEEKRFKLD HSVSSTNGHRWQIPQDWLGTDQDNLDLANV NLMLELLVQKKKQLEAESHAAQLQILMEFLK VARRNKREQLEQIQKELSVLEEDIKRVEEMS GLYSPVSEDSTVPQFEAPSPSHSSIIDSTEYSQP PGFSGSSQTKKQPWYNSTLASRRKRLTAHFE DLEQCYFSTRMSRISDDSRTASQLDEFQECLIS KF\TRYNSVRPLATLSYASDLYNGSQYKSLV FEFDRDCDYFAIAGVTKKIKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHLAFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKPYYCLRSFKGHIN EKNFVGLASNGDYIACGSENNSLYLYYKGLS KTLLTTKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGTTKVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPFTS ASQVAEITSVRHITWLIFCILGQMGFHHVGE QAGLELLTSWPALPSQSAGIIGMSPHAWPP 598 1948 A 4524 I 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF	594	1944	A	4507	1327	647	
CTYGKPVTFDCAVKPSVTCVDQDFKSQKNFII NMTCRFCWQLPETDYECTNSTSCMTVSCPRQ RYPANCTVR\DFIVHCLGNRTFPKMLYCNWT GGYKWVYGLWLLRHHPRWGLGADRF\YLGP VAGTASGKLFSFGGLGIWTLIDVLLIGVGYVG PADGSLYI 595 1945 A 4512 533 264 FFFKMESYSVARLECSGAISAPCNLHLLGSNN SPASASRV/AGNIGARHHTQQIFVILVQMRVH YVGQDGLDLL/NLMIHPPRSPKVLGLQA HSVSSTNGHRWQIFQDWLGTDQDNLDLANV NLMLELLVQKKKQLEAESHAAQLQILMFFLK VARRNKEQLEQIQKELSVLEEDIKRVEEMS GLYSPVSEDSTVPQFEAPSPSHSSIIDSTEYSQP PGFSGSSQTKKQPWYNSTLASRRKRLTAHFE DLEQCYFSTRMSRISDDSRTASQLDEFQEC\LS KF\TRYNSVRPL\ATLS\YASDL\YNGSQYKSLV FFFDDCDYFAIAGVTKKIKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSYYHKNLLASS DYEGTVIL WDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHLAFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKP\YCLRSFKGHIN EKNFVGLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGTNKVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPTTS ASQVAEITSVRHITVLIFCLLGQMGFHHVGE QAGLELLTSVRHITVLIFCLLGQMGFHHVGE QAGLELLTSVRHITVLIFCLLGQMGFHHVGE QAGLELTSVRHITVLIFCLLGQMGFHHVGE QAGLELTSVRHITVLIFCLLGQMGFHHVGE QAGLELTSVRHITVLIFCLLGQMGFHHVGE							
SPR SPR							
RYPANCTVR\DHYHCLGNRTFPKML\YCNWT\ GGYKWYGLWLLRHHPWGLGADRF\YLGP\ VAGTASGKLFSFGGLGIWTLIDVLLIGVGYVG\ PADGSLYI\ 595	1 1						
GGYKWYYGLWLLRHIPRWGLGADRF\YLGP	i						RYPANCTVR\DHVHCLGNRTFPKMLYCNWT
VAGTASGKLFSFGGLGIWTLIDVLLIGVGYVG							
PADGSLYI		l					
1945	i 1						
SPASASRV/AGNIGARHHTQQIFVLLVQMRVH YVGQDGLDLL/NLMIHPPRSPKVLGLQA 596 1946 A 4513 1674 HASDHLYPNFLVNFLILKQKQRFEEKRFKLD HSVSSTNGHRWQIFQDWLGTDQDNLDLANV NLMLELLVQKKKQLEAESHAAQLQILMEFLK VARNKREQLEQIQKELSVLEEDIKRVEEMS GLYSPVSEDSTVPQFEAPSPSHSSIIDSTEYSQP PGFSGSSQTKKQPWYNSTLASRRKRLTAHFE DLEQCYFSTRMSRISDDSRTASQLDEFQECLS KFTRYNSVRPLATLSYASDLYNGSQYKSLV FEFDRDCDYFAIAGVTKKIKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCVKFSPSSRYHLAFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKPYCLRSFKGHIN EKNFV\GLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGTNKVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCILGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 I 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF	595	1945	Ā	4512	533	264	
996 1946 A 4513 3 1674 HASDHLYPNFLVNFLILKQKQRFEEKRFKLD HSVSSTNGHRWQIFQDWLGTDQDNLDLANV NLMLELLVQKKKQLEAESHAAQLQILMEFLK VARRNKREQLEQIQKELSVLEEDIKRVEEMS GLYSPVSEDSTVPQFEAPSPSHSSIIDSTEYSQP PGFSGSSQTKKQPWYNSTLASRRKRLTAHFE DLEQCYFSTRMSRISDDSRTASQLDEFQECCLS KF\TRYNSVRPL\ATLSYASDLYNGSQYKSLV FFFDRDCDYFAIAGVTKKIKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHI.AFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKPYCLRSFKGHIN EKNFV\GLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGTIKVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCRLGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 I 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF					333	20.	
1946 A 4513 3 1674 HASDHLYPNFI.VNELILKQKQRFEEKRFKLD HSVSSTNGHRWQIFQDWLGTDQDNLDLANV NLMLELLVQKKKQLEAESHAAQLQILMEFLK VARRNKREQLEQIQKELSVLEEDIKVEEMS GLYSPVSEDSTVPQFEAPSPSHSSIIDSTEYSQP PGFSGSSQTKKQPWYNSTLASRRKRLTAHFE DLEQCYFSTRMSRISDDSRTASQLDEFQEC\LS KF\TRYNSVRPL\ATLSYASDLYNGSQYKSLV FEFDRDCDYFAIAGVTKKIKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHI.AFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKP\YCLRSFKGHIN EKNFV\GLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPPGESNVLIAANS\QGTIKVLELV CWRALPDGESNVLIAANS\QGTIKVLELV GWAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 I 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF							
HSVSSTNGHRWQIFQDWLGTDQDNLDLANV NLMLELLVQKKKQLEAESHAAQLQILMEFLK VARRNKREQLEQIQKELSVLEEDIKRVEEMS GLYSPVSEDSTVPQFEAPSPSHSSIIDSTEYSQP PGFSGSSQTKKQPWYNSTLASRRKRLTAHFE DLEQCYFSTRMSRISDDSRTASQLDEFQECLLS KF\TRYNSVRPL\ATLSYASDLYNGSQYKSLV FEFDRDCDYFAIAGVTKKIKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHLAFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKP\YCLRSFKGHIN EKNFVGLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGTI\KVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCI\LGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 I 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF	596	1946	Α	4513	3	1674	HASDHI VPNEL WIEL II KOKOBEEEKBERI D
NLMLELLVQKKKQLEAESHAAQLQILMEFLK VARRNKREQLEQIQKELSVLEEDIKRVEEMS GLYSPVSEDSTVPQFEAPSPSHSSIIDSTEYSQP PGFSGSSQTKKQPWYNSTLASRRKRLTAHFE DLEQCYFSTRMSRISDDSRTASQLDEFQEC\LS KF\TRYNSVRPL\ATLSYASDLYNGSQYKSLV FEFDRDCDYFAIAGVTKKIKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHLAFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKPYCLRSFKGHIN EKNFV\GLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGTNKVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCILGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 I 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF		17.0	**	1313	3	1074	HSVSSTNGUDWOIFODWI GTDODNI DI ANIV
VARRNKREQLEQIQKELSVLEEDIKRVEEMS GLYSPVSEDSTVPQFEAPSPSHSSIIDSTEYSQP PGFSGSSQTKKQPWYNSTLASRRKRLTAHFE DLEQCYFSTRMSRISDDSRTASQLDEFQEC\LS KF\TRYNSVRPL\ATLSYASDLYNGSQYKSLV FEFDRDCDYFAIAGVTKKIKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHLAFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKP\YCLRSFKGHIN EKNFV\GLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPPDGESNVLIAANS\QGTNKVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCILGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 I 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF							MI MI ELI VOVVVOI EAEGUAAOLOU MEELV
GLYSPVSEDSTVPQFEAPSPSHSSIIDSTEYSQP PGFSGSSQTKKQPWYNSTLASRRKRLTAHFE DLEQCYFSTRMSRISDDSRTASQLDEFQEC\LS KF\TRYNSVRPL\ATLSYASDLYNGSQYKSLV FEFDRDCDYFAIAGVTKKIKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHI.AFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKPYCLRSFKGHIN EKNFV\GLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGT\KVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCILGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 I 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF							
PGFSGSSQTKKQPWYNSTLASRRKRLTAHFE DLEQCYFSTRMSRISDDSRTASQLDEFQEC\LS KF\TRYNSVRPL\ATLSYASDLYNGSQYKSLV FEFDRDCDYFAIAGVTKKIKVYEYDTVIQDA VDHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHI.AFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKP\YCLRSFKGHIN EKNFV\GLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGT\KVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCILGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 I 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF							
DLEQCYFSTRMSRISDDSRTASQLDEFQEC\LS KF\TRYNSVRPL\ATLSYASDLYNGSQYKSLV FEFDRDCDYFAIAGVTKKIKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHI.AFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKP\YCLRSFKGHIN EKNFV\GLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGTRKVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCRLGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 I 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF							
KF\TRYNSVRPL\ATLSYASDLYNGSQY\KSLV FEFDRDCDYFAIAGVTKKIKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHI.AFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKP\YCLRSFKGHIN EKNFV\GLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGT\KVLELV CWRALPDGESNVLIAANS\QGT\KVLELV ASQVAEITSVRHHTWLIFC\LGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVF\\CRUPGAYFFSF]				
FEFDRDCDYFAIAGVTKKIKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHI.AFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKP\YCLRSFKGHIN EKNFV\GLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGT\KVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCRLGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 I 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF							
VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHI.AFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKP\YCLRSFKGHIN EKNFV\GLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGTNKVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCRLGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 I 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF							ACTICINOVELLATES YASDLYNGSQYKSLV
DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHI.AFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKP\YCLRSFKGHIN EKNFV\GLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPPDGESNVLIAANS\QGTNKVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCI\LGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 I 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF	1	- 1					
DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHI.AFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKPYYCLRSFKGHIN EKNFV\GLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGTNKVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCILGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 I 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF		ļ					
ASIEAKANVCCVKFSPSSRYHI.AFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKPYYCLRSFKGHIN EKNFV\GLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGT\KVLELV 597 1947 A 4518 536 824 RSLALSPGIECSGMISAHCNLHLLGSSDPTS ASQVAEITSVRHHTWLIFCILGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 I 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF							
HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKP\YCLRSFKGHIN EKNFV\GLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGTI\KVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCILGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 I 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF		ļ	l				
EEIVSASTDSQLKLWNVGKP\YCLRSFKGHIN EKNFV\GLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGTI\KVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCI\LGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 I 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF		İ	ļ			į	
EKNFV\GLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGTI\KVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCRLGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 I 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF		ļ	İ			j	
KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGTI\KVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCRLGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 I 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF		1					,
597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCRLGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF				:		ľ	
597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCRLGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF		1					
ASQVAEITSVRHHTWLIFCRLGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 I 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF							CWRALPDGESNVLIAANS\QGTI\KVLELV
598 1948 A 4524 I 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF	597	1947	A	4518	536	824	
598 1948 A 4524 I 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF]	İ	ļ	}		ASQVAEITSVRHHTWLIFCI\LGQMGFHHVGE
598 1948 A 4524 I 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF					}		
	598	1948	A	4524	1	384	
					ł		

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ . ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion RRREMQSQSVMLALRRGDAVWLLSHDHDG
						YGAYSNHGKYITFSGFLVYPDLAPAAPPGLG ASELL
599	1949	A	4526	366	776	MGQPAPYAEGPIQGGDAGELCKCDFLVFTSP NPEAVCEAGTPAMFQTAWRQMESCSI/AQAG VQWRDPGSLHPPPLGFKRFSCLSLPSSWDYK HAPPHPANFCIFSRDQVSPCWPGWSRSLDLVI PPPWLPKVLGLQA
600	1950	A	4529	776	334	FFFETESCYVAQAGVQWCDLCSLQAPPPG\SS DPPASASRVAGTTGARHHTQLIFVFLVETGFH \MLARDGLKLLTSSDPPASASQSSWDYRREPP RLANFFVFLVETGSRYVAQAGVQWLFTGAIP LLISTGVLTCSVSDLGRFTPP
601	1951	A	4533	1460	403	HEVQESIHFLESEFSRGISDNYTLALITYALSS VGSPKAKEALNMLTWRAEQEGGMQFWVSSE SKLSDSWQPRSLDIEVAAYALLSHFLQFQTSE GIPIMRWLSRQRNSLGGFASTQDTTVALKALS EFAALMNTERTNIQVTVTGPSSPSPVKFLIDT HNRLLQTAELADGTANGSVISIANGFGFAI CQLNVVYNVKASGSSRRRSIQNQEAFDLDV AVKENKDDLNHVDLNVCTSFSGPGRSGMAL MEVNLLSGFMVPSEAISLSETVKKVEYDHGK LNLYLDSVNETQFCVNIPAVRNFKVSNTQDA SVSIVDYYEPRRQAVRSYNSEVKLSSCDLCSD VQRLPSL
602	1952	A	4540	1963	295	MRAPGRPALRPLPLPPLLLLLLSSPWGRAVPC VSGGLPKPANITFLSINMKNVLQWTPPEGLQG VKVTYTVQYFIYGQKKWLNKSECRNINRTYC DLSAETSDYEHQYYAKVKAIWGTKCSKWAE SGRFYPFLETQIGPPEVALTTDEKSISVVLTAP
403	1052					EKWKRNPEDLPVSMQQIYSNLKYNVSVLNT KSNRTWSQCVTNHTLVLTWLEPNTLYCVHV ESFVPGPPRRAQPSEKQCARTLKDQSSEFKAK IIFWYVLPISITVFLFSVMGYSIYRYIHVGKEK HP\ANLILIYG\NEFDKRFFVPA\EKIV\INFI\TL NIS\DDSKISHQDMSLLGKSSDVSSLNDPQPSG NLRPPQEEEEVKHLGYASHLMEIFCDSEENT\ EGTSFTQQESLSRTIPPDKTVIEYEYDVRTTDI CAGPEEQELSLQEEVSTQGTLLESQAALAVL GPQTLQYSYTPQLQDLDPLAQEHTDSEEGPEE EPSTTLVDWDPQTGRLCIPSLSSFDQDSEGCE PSEGDGLGEEGLLSRLYEEPAPDRPPGENETY LMQFMEEWGLYVQMEN
603	.1953	A	4543	3	600	YSAVEFVEQASGISDWWNPALRKRMLSDSGL GMIAPYYEDSDLKDLSHSRVLQSPVSSEDHAI LQAVIAGDLMKLIESYKNGGSLLIQGPDHCSL LHYAAETGNGEIVKYILDHGPSELLDMADSE TGETALHKAACQRNRAVCQLLVDAGASLRK\ TDSKGKTPQERAQQA\GDPDLAAYYTIESRQN YKVIGHEDLETAV
604	1954	A	4548	3	938	QDNKVQNGSLHQKDTVHDNDFEPYLTGQAN QSNSYPSMSDPYLSSYYPPSIGFPYSLNEAPW STAGDPPIPYLTTYGQLSNGDHHFMHDAVFG QPGGLGNNIYQHRFNFFPENPAFSAWGTSGS QGQQTQSSAYGSSYTYPPSSLGGTVVDGQPG FHSDTLSKAPGMNSLEQGMVGLKIGDVSSSA VKTVGSVVSSVALTGVLSGNGGTNVNMPVS KPTSWAAIASKPAKPQPKMKTKSGPVMGGG LPPPPIKHNMDIGTWDNKGPVPKAPVPQQAP

	l and in	136	1.000			
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	j	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ľ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		ļ		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	1		Ì	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	ļ			peptide		/=possible nucleotide deletion, \=possible
	ļ	<u> </u>		sequence		nucleotide insertion
		1				SPQAAPQPQQVAQPLPAQPPALAQPQYQSPQ
	1000	ļ.,	<u> </u>			QPPQ
605	1955	Α	4553	2	2304	ILLQEKRNCLLMQLEEATRLTSYLQSQLKSLC
						ASTLTVSSGSSRGSLASSRGSLASSRGSLSSVS
		-	;		!	FTDIYGLPQYEKPDAEGSQLLRFDLIPFDSLGR
						DAPFSEPPGPSGFHKQRRSLDTPQSLASLSSRS
			1			SLSSLSPPSSPLDTPFLPASRDSPLAQLADSCE
	1	1	!			GPGLGALDRLRAHASAMGDEDLPGMAALQP
						HGVPGDGEGPHERGPPPASAPVGGTVTLRED
					į	SAKRLERRARRISACLSDYSLASDSGVFEPLT
	I	1				KRNEDAEEPAYGDTASNGDPQIHVGLLRDSG
	1				ļ	SECLLVHVLQLKNPAGLAVKEDCKVHIRVYL
		ŀ				PPLDSGTPNTYCSKALEFQVPLVFNEVFRIPV
	1	1	1			HSSALTLKSLQLYVCSVTPQLQEELLGIAQIN
						LADYDSLSEMQLRWHSVQVFTS\LNHQGRGR
						LGVQERAPPGTLHTPSPSPA/STDAVTVLLAR
	1	l				TTAQLQAVERELAEERAKLEYTEEEVLEMER
						KEEQAEAISERSWQADSVDSGCSNCTQTSPPY
	1		1			PEPCCMGIDSILGHPFAAQAGPYSPEKFQPSPL
	i					KVDKETNTEDLFLEEAASLVKERPSRRARGSP
	1	1]			FVRSGTIVRSQTFSPGARSQYVCRLYRSDSDS
	1	ì				STLPRKSPFVRNTLERRTLRYKQSCRSSLAEL
						MARTSLDLELDLQASRTRQRQLNEELCALRE
	1	1				LRQRLEDAQLRGQTDLPPWVLRDERLRGLLR
	l]	l'i			EAERQTRQTKLDYRHEQAAEKMLKKASKEI
						YQLRGQSHKEPIQVQTFREKIAFFTRPRINIPPL
		l	1 1		l	PADDV
606	1956	Α	4555	3429	776	PGSGPGPAPFLAPVAAPVGGISFHLQIGLSREP
	1 .	ļ	l ;			VLLLQDSSGDYSLAHVREMACSIVDQKFPEC
	İ	1	1 1	i		GFYGMYDKILLFRHDPTSENILQLVKAASDIQ
			1			EGDLIEVVLSASATFEDFQIRPHALFVHSYRA
			}			PAFCDHCGEMLWGLV\RQGLKCEGCGLNYH
			1 1	(KRCAFKIPNNCSGVRRRRLSNVSLTGVSTIRT
						SSAELSTSAPDEPLLQKSPSESFIGREKRSNSQ
			1			SYIGRPIHLDKILMSKVKVPHTFVIHSYTRPTV
				1		CQYCKKLLKGLFRQGLQCKDCRFNCHKRCA
]	}		PKVPNNCLGEVTINGDLLSPGAESDVVMEEG
			1			SDDNDSERNSGLMDDMEEAMVQDAEMAMA
						ECQNDSGEMQDPDPDHEDANRTISPSTSNNIP
		1		}		LMRVVQSVKHTKRKSSTVMKEGWMVHYTS
				!		KDTLRKRHYWRLDSKCITLFONDTGSRYYKE
				Ì		IPLSEILSLEPVKTSALIPNGANPHCFEITTANV
	[]				ľ	VYYVGENVVNPSSPSPNNSVLTSGVGADVAR
						MWEIAIQHALMPVIPKGSSVGTGTNLHRDISV
]					SISVSNCQIQENVDISTVYQIFPDEVLGSGQFGI
1				1	ĺ	VYGGKHRKTGRDVAIKIIDKLRFPTKOESOLR
				1		NEVAILQNLHHPGVVNLECMFETPERVFVVM
]]	1		EKLHGDMLEMILSSEKGRLPEHITKFLITOILV
ļ			[1	ſ	ALRHLHFKNIVHCDLKPENVLLASADPFPOV
1	j			İ		KLCDFGFARIIGEKSFRRSVVGTPAYLAPEVL
ļ			J i			
ļ	1 1			[RNKGYNRSLDMWSVGVIIYVSLSGTFPFNED
						EDIHDQIQNAAFMYPPNPWKEISHEAIDLINN
ì				1		LLQVKMRKRYSVDKTLSHPWLQDYQTWLDL
1] [I	ı	DEL COMMODDA MINISTER DE SANCIONA DE CARACTER DE COMPONIONE DE COMPONION
,						RELECKIGERYITHESDDLRWEKYAGEQGLQ
						YPTHLINPSASHSDTPETEETEMKALGERVSIL
607	1957	A	4563	1	4499	YPTHLINPSASHSDTPETEETEMKALGERVSIL SRPWWLRASERPSAPSAMAKRSRGPGRRCLL
607	1957	A	4563	1	4499	YPTHLINPSASHSDTPETEETEMKALGERVSIL SRPWWLRASERPSAPSAMAKRSRGPGRRCLL ALVLFCAWGTLAVVAQKPGAGCPSRCLCFRT
607	1957	A	4563	1	4499	YPTHLINPSASHSDTPETEETEMKALGERVSIL SRPWWLRASERPSAPSAMAKRSRGPGRRCLL ALVLFCAWGTLAVVAQKPGAGCPSRCLCFRT TVRCMHLLLEAVPAVAPQTSILDLRFNRIREI
607	1957	A	4563	1	4499	YPTHLINPSASHSDTPETEETEMKALGERVSIL SRPWWLRASERPSAPSAMAKRSRGPGRRCLL ALVLFCAWGTLAVVAQKPGAGCPSRCLCFRT

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion HFNQIETLDPDSFQHLPKLERLFLHNNRITHL VPGTFNHLESMKRLRLDSNTLHCDCEILWLA DLLKTYAESGNAQAAAICEYPRRIQGRSVATI TPEELNCERPRITSEPQDADVTSGNTVYFTCR AEGNPKPEIIWLRNNNELSMKTDSRLNLLDD GTLMIQNTQETDQGIYQCMAKNVAGEVKTQ EVTLRYFGSPARPTFVIQPQNTEVLVGSSVTL ECSATGHPPPRISWTRGDRTPLPVDPRVNITPS GGLYIQNVVQGDSGEYACSATNNIDSVHATA FIIVQALPQFTVTPQDRVVIEGQTVDFQCEAK GNPPPVIAWTKGGSQLSVDRRHLVLSSGTLRI SGVALHDQGQYECQAVNIIGSQKVVAHLTVQ PRVTPVFASIPSDTTVEVGANVQLPCSSQGEP EPAITWNKDGVQVTESGKFHISPEGFLTINDV GPADAGRYECVARNTIGSASVSMVLSVNVPD VSRNGDPFVATSIVEAIATVDRAINSTRTHLF DSRPRSPNDLLALFRYPRDPYTVEQARAGEIF ERTLQLIQEHVQHGLMVDLNGTSYHYNDLVS PQYLNLIANLSGCTAHRRVNNCSDMCFHQKY RTHDGTCNNLQHPMWGASLTAFERLLKSVY ENGFNTPRGINPHRLYNGHALPMPRLVSTTLI GTETVTPDEQFTHMLMQWGQFLDHDLDSTV VALSQARFSDGQHCSNVCSNDPPCFSVMIPPN DSRARSGARCMFFVRSSPVCGSGMTSLLMNS VYPREQINQLTSYIDASNVYGSTEHEARSIRD LASHRGLLRQGIVQRSGKPLLPFATGPPTECM RDENESPIPCFLAGDHRANEQLGLTSMHTLW FREHNRIATELLKLNPHWDGDTIYYETRKIVG AEIQHITYQHWLPKILGEVGMRTLGEYHGYD
608	1958	A	4566	354	1135	PGINAGIFNAFAT\AAFRGHTLVNPLLLPGLD ENEQPIAQDHI.PLHKAFFSPFRIVAISGGIDPLL RGLFGVAGKMRVPSQLLNTELTERLFSMAHT VALDLAANIQRGRDHGIPPYHDYRVYCNLS AAHTFEDLKNEIKNPEIREKLKRLYGSTLNID LFPALVVEDLVPGSRLGPTLMCLLSTOFKRLR DGDRLWYENPGVFSPAQLTQIKQTSLARILCD NADNITRVQSDVFRVAEFPHGYGSCDEIPRVD LRVWQDCCEDCRTRGQFNAFSYHFRGRRSLE FSYQEDKPTKKTRPRKIPSVGRQGEHLSNSTS AIFSTRSDASGITNDFQRVCSWEMQKTITDLR TQIKKLESRLSTTECVDAGGESHANNTKWK KDACTICECKDGQVTCFVEACPPATCAVPVNI PGACCPVCLQKRAEEKP FSFLC/GVSGRLGLDSEEDYYTPQKVDVPKAL IIVAVQCGCDGTFLLTQSGKVLACGLNEFNKL GLNQCMSGINHEAYHEVPYTTSFTLAKQLSF YKIRTIAPGKTHTAAIDERGRLLTFGCNKCGQ LGVGNYKKRLGINLLGGPLGGKQVIRVSCGD EFTIAATDDNHIFAWGNGGNGRLAMTPTERP HGSDICTSWPRPIFGSLHHVPDLSCRGWHTILI VEKVLNSKTIRSNSSGLSIGTVFQSSSPGGGE GGPDAW FFFFETESRSVAQAGVQWRDLGSLQAPPPGFT PFSCLSLPSSWDYRRPPLRPANFFVFLVETGF HRFSRDGLDLLT/S/GDPPASASQSAGITGVSH RARPRINLRNVIYSFAVTYCLNYISLAMSSTL
610	1960	A	4570	697	467	KLSFHVLSGS ECRGVISAH\CCTLCLPSSSDSASAF\RVARTT GTCDYAQLIFAFLVEMGFHHVGQDGLHILL'N LVIRPPRPPKVLGLQA

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in USSN	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide seq-	seq- uence		055N 09/496	location correspondi	to last amino	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
uence	licited		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
			***	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
İ				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
	l	1		sequence		nucleotide insertion
611	1961	Α	4571	25	1396	ADPHTTVIRFFPAASATKRVLPPVLRVSSPRT
1						WNPNVPESPRIPAPRLPKRMSGAPTAGAALM
i		ł	1	1	}	LCAATAVLLSAQGGPVQSKSPRFASWDEMN
						VLAHGLLQLGQG\CANT\GAHPQSAERAGA\R
	ļ					LSACGSACQGTEGSTDLPLAPESRVDPEVLHS
ļ	ļ	l	1	ļ]	LQTQLKAQNSRIQQLFHKVAQQQRHLEKQHL
						RIQHLQSQFGLLDHKHLDHEVAKPARRKRLP EMAQPVDPAHNVSRLHRLPRDCQELFQVGER
				İ		QSGLFEIQPQGSPPFLVNCKMTSDGGWTVIQR
						RHDGSVDFNRPWEAYKAGFGDPHGEFWLGL
	}	1	1			EKVHSITGDRNSRLAVQLRDWDGNAELLQFS
1		[1			VHLGGEDTAYSLQLTAPVAGQLGATTVPPSG
						LSVPFSTWDQDHDLRRDKNCAKSLSGGWWF
						GTCSHSNLNGQYFRSIPQQRQKLKKGIFWKT
		İ				WRGRYYPLQATTMLIQPMAAEAAS
612	1962	A	4575	162	3	FFFETESRSVAQAGVQWRDLSSLQPPPPG\SR
(12	1060	ļ		600		GSPASASPVAGITGTRHHRTRG
613	1963	Α	4584	687	321	PLAQRRPFLWVTVKTNGHIWGSSTYPHFWGS
	1	ļ			ļ	SNS/PASASQVAGIPNARHQARIIFVFLVEPRF
						HHVGRAGLGFL/NLAICLPQHPKVLGLQACN LNIKPHPAHKYISMIQFNVHFMCMSVHIYI
614	1964	A	4589	727	299	PGSAQSAQRGRGRRRARAGSATQITMYSFMG
011	1701	ļ ' '	430)	, 2,	2,,,	GGLFCAWVGTILLVVAMATDHWMQYRLSGS
						FAHQGLWRYCLGNKCYLQTDSIAYWNATRA
						FMILSALCAISGIIMGIMAF/GWVAVLMTFFA
						GIFYMCAYRVHECRRLSTPR
615	1965	A	4590	2	414	TILPEKIQAWAQKQCPQSGEEAVALVVHLEK
						ETGRLRQQVSSPVHREKHSPLGAAWEVADFQ
						PEQVETQPRAVSREEPGSLHSGHQEQLNRKR
						ERRPLPKNARPSPWVPALADEWNTLHQEVTT
616	1966	A	4592	773	488	TRLPAGSQEPVKD DFALVAQAGVQWHNLGSPQPLPPGFKRFSCL
0.0	1300	А	4372	113	400	SLPSSWEYRCVPP/RLANFVFLVEMGFLHVGO
						AGLELPTSGDPPALASQSAGITGVTTVPSGPG
617	1967	В	4595	84	478	XRHGLREPLLERRCAAASSFQHSSSLGRELPY
				_		DPVDTEGFGEGGDMQERFLFPEYILDPEPQPT
						REKQLQELQQQQEEEERQRQQRREERRQQNL
						RARSREHPVVGHPDPALPPSGVNCSGCGAEL
						HCQDAR*
618	1968	Α	4596	2945	1188	ARSRNSARGVYGMCVDTLFLCFLEDLERNDG
	}					SAERPYFMCSTLKKPLARRCFPAIHAYKGVL
				ļ	ļ	MVGNETTYEDGHGSRKNITDLVEGAKKANG
				ŀ		VLEARQLAMRIFEDYTVSWYWIIIGLVIAMA MSI I SIII I HI I AGIMGWZMIMENSEI CYPIE
			1			MSLLSIILLHLLAGIMGWVMIIMEI\SELGYRIF HCYMEYSRLRGEAGSDVSLVDLGFOTDFRV
						YLHLRQTWLAFMILSILEVIIILLLIFLRKRILI
						AIALIKEASRAVGYVMCSLLYPLVTFFLLCLCI
	ļ	' I				AYWASTAVFLSTSNEAVYKIFDDSPCPFTAKT
						CNPETFPSSNESRQCPNARCQFAFYGGESGYH
						RALLGLQIFNAFMFFWLANFVLALGQVTLAG
		ľ			}	AFASYYWALRKPDDLPAFPLFSAFGRALRYH
					ľ	TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN
					ļ	KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM
						IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV
		J		j	j	TDFLFLLGKLLIVGSVGILAFFFFTHRIRIVQDT
						APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC
			ļ			VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES
619	1969	A	4601	2	357	RTSVEPYILGEF/RKLSNNTKVVKTEYKATEY
717	1707	43	1001	_	100	KISTEL HEGETAKESININIK VALETKALET

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	l lou	in in	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	}	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	ļ		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	j	,	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		İ		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon.
1	1	1	1	peptide]	/=possible nucleotide deletion, \=possible
— —	<u> </u>	ļ	ļ	sequence		nucleotide insertion
]]	GLAYGHFSYEFSNHRDVVVDLQGWVTGNGK
-			1		i	GLIYLTDPQIHSVDQKVFTTNFGKRGIFYFFN
620	1970	A	4606	1	2415	NQHVECNEICHRLSLTRPSMEKPCKS
***	1 .,,,	^	1 4000	,	2413	MERLWGLFQRAQQLSPRSSQTVYQRVEGPR KGHLEEEEEDGEEGAETLAHFCPMELRGPEP
	1	1				LGSRPRQPNLIPWAAAGRRAAPYLVLTALLIF
-			ł			TGAFLLGYVAFRGSCQACGDSVLVVSEDVN
]			1		ļ	YEPDLDFHQGRLYWSDLQAMFLQFLGEGRL
		1				EDTIRQTSLRERVAGSAGMAALTQDIRAALS
						RQKLDHVWTDTHYVGLQFPDPAHPNTLHWV
		ĺ				DEAGKVGEQLPLEDPDVYCPYSAIGNVTGEL
						VYAHYGRPEDLQDLRARGVDPVGRLLLVRV
						GVISFAQKVTNAQDFGAQGVLIYPEPADFSQ
	İ			·	,	DPPKPSLSSQQAVYGHVHLGTGDPYTPGFPSF
		{				NQTQFPPVASSGLPSIPAQPISADIASRLLRKL
	ļ	1				KGPVAPQEWQGSLLGSPYHLGPGPRLRLVVN
						NHRTSTPINNIFGCIEGRSEPDHYVVIGAQRDA WGPGAAKSAVGTAILLELVRTFSSMVSNGFR
						PRRSLLFISWDGGDFGSVGSTEWLEGYLSVL
					•	HLKAVVYVSLDNAVLGDDKFHAKTSPLLTSL
		,				IESVLKQVDSPNHSGQTLYEQVVFTN\PSWD\
1] [AEVIRPLPM\DSSAY\SFTAFVGVPAVEFSFME\
			1			DDQ\AYPFLHTKEDTYENLHKVLQGRLPAVA
		[[[QAVAQLAGQLLIRLSHDRLLPLDFGRYGDVV
1			1	Ţ		LRHIGNLNEFSGDLKARGLTLQWVYSARGDY
	1					IRAAEKLRQEIYSSEERDERLTRMYNVRIMRV
		İ				EFYFLSQYVSPADSPFRHIFMGRGDHTLGALL
			ĺĺ			DHLRLLRSNSSGTPGATSSTGFQ\ESRFRRQL\ ALL\TWDACKGAANALSGDVWNIDNNF
621	1971	A	4610	793	334	ISRVDDFVGSGIANVIIAVAIFSIPAFARLVRG\
	L					NTLVLKQQTFIESARSIGASDMTVLLRHILPGT
					·	GSSIVVFFTMRIGTSIISAASLSFLGLGAQPPTP
				į		EWGAMLNEARADMVIAPHVAVFPALAIFLTV
(00						LAFNLLGDGLRDALDPKIKG
622	1972	A	4614	2	820	LVYVMIAIFCIASAMSLYNCLAALIHKIPYGQ
			i			CTIACRGKNMEVRLIFLSGLCIAVAVVWAVF
					ĺ	RNEDRWAWILQDILGIAFCLNLIKTLKLPNFK
				1		SCVILLGLLLYDVFFVFITFITKNGESIMVEL
				Ì		AAGPFGNNEKNDGNLVEATGQPSAPHEKLPV VIRVPKLIYFSVMSVCLMPVSILGFGDIIVPGL
				l		LIAYCRRFDVQTGSSYIYYVSV\TVAYAIGMIL
[ĺ	· [i	ľ	TFVVLG\LMKKGQPALLYLVPCTLITA/CQFV
				!	1	AWETVREMKKFWERVTS
623	1973	A	4619	17	691	TLVSVVEFVRRADLTREDLAPSSVDSGQAGF
					j	GGCCESGLPNTMPSAFSVSSFPVSIPAVLTOT
			ł		l	DWTEPWLMGLATFHALCVLLTCLSSRSYRLO
	}	-	1	i	ł	IGHFLCLVILVYCAEYINEAAAMNWRLFSKY
					ŀ	QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW
	l	I	1	}	ł	KTLNVMTDLKNAQERRKEKKRRKED*GAA
]	ĺ	1			AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG
624	1974	A	4622	164	668	LKNRCFI VSCVTALOSIMNOPESANDREDI CAVOCOALI
'	****		.022	101	000	VSCYTALQSIMNQPESANDPEPLCAVCGQAH SLEENHFYSYPEEVDDDLICHICLQALLDPLD
	- 1	l				TPCGHTYCTLCLTNFLVEKDFCPMDRKPLVL
		·	ļ			QHCKKSSILVNKLLNKLLVTCPFREHCTQVL
		- 1	į			QRCDLEHHFQTSQAWGTHL*SQLLGRLRQED
						CLSPGVHHCSEV
625	1975	Α	4625	474	473	CFLSPSPLLPPLLLSSSSSPSFPLPPPPTLLPSTLP
Ll						PPLLIPSS*LSP

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	l	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq- uence		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq- uence	uence		09/496 914	correspondi ng to first	to last amino acid residue	M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
uence			914	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
į		[residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ŀ		peptide	sequence	/=possible nucleotide deletion, \=possible
ŀ				sequence		nucleotide insertion
626	1976	A	4629	249	3	KLKGNECFCYHCNVCIFLMIKK*GLFLC*IYFI
		ļ	i		_	LFFET*SHSFTRLECSGTISAHCSLQLQGSSNSP
		1		Í	[ASASQVAGIAGTHH
627	1977	A	4635	1	301	FFFFETKPFFAPQAGGQGPSRGSLNPLPTGLK
						QFSGLTLSRSGNNGPRPPPRVNFGILRGNGVP
						PGGAG*PRPPDLRGPPGLAPPQGGNNGGDPP
						ARAYL
628	1978	Α	4648	1357	782	KLFSSQRLFGPHIQAINPSFLLLSFFPS*LLAMR
						TVGNNAFILVFLVYRIVLLLF*HV*PAYFQPSK
						NKTAKINCN*RPFLFLVCYLL*AELHIGIFIANF
			ļ			YDCIPNKLNEHLWPKLLQSLIFHVDFCGFLHK
						VFYICFTEFLLFLYFL*LFIIKV\$CSII*CSTICVF
620	1070		1660	10	000	SYKSFAVIIFFVDNTRFFSFGF
629	1979	Α	4660	18	999	HHELHTLELLQNPKEVLTRSEIQDVNYSLEAV
		1]	KVKTVCQIPLMKEMLKRFQVAVNLAEDTAH PKLVFSQEGRYVKNTASASSWPVFSSAWNYF
						AGWRNPQKTAFVERFQHLSCVLGKNVFTSG
						KHYWEVESRDSLEVAVGVCREDVMGITDRS
				'		KMSPDVGIWAIYWSAAGYWPLIGFPGTPTOO
1						EPALHRVGVYLDRGTGNVSFYSAVDGVHLH
						TFSCSSVSRLRPFFWLSPLASLVIPPVTDRK*G
]			FSSPDQNSFPVVQLRDTHPWALFCPSCLYPG
1 1			!			WSIFWVSLTVPFGICPLCASQEAVPWEVGLA
1					•	NGDGTGNFPRRFWEIFL
630	1980	Α	4669	2	358	FFFFFETESHSVAQAGMQWRNLGSLPAPPPGF
1 1			ļ i			TPFFCLSLLNGWDYRRPPPHLANFFVLLVETG
1 1						FHDVGQDGLDLLTS*STPSASQSAEITGVSHC
						TRLKKIRFAKGHVEFFFESHVE
631	1981	Α	4674	953	614	TPIRGTDDEHEECTVQEYSAGKNTCLRPGAV
						AHTCNPCTLGGRGRWIT*GSGVQDQPGPTWQ
						NPVFLERRPRALHSSPGLTTQRILWAQGLWV
632	1982	A	4670	24	244	GAGSTGCSRGPRGEGVFREG
032	1982	Α	4678	34	314	RSTHASGMISPSFGFMGHLLRLEFEILPSTPNP
						*LPSYQGEAAGSSLISHLQTFSPDLKGVYCTFP ASGLAPVPTHWTVSELSRSPVATATFC
633	1983	Α	4696	1	1365	RTLGMEGERRASQAPSSGLPAGGANGESPGG
""	1703	71	0,00	*	1303	GAPFPGSSGSSALLQAEVLDLDEDEDDLEVFS
	ļ					KDASLMDMNSFSPMMPTSPLSMINQIKFEDEP
						DLKDLFITVDEPESHVTTIETFITYRITKTSRG
				ļ		EFDSSEFEVRRRYQDFLWLKGKLEEAHPTLII
						PPLPEKFIVKGMVERFNDDFIETRRKALHKFL
					İ	NRIADHPTLTFNEDFKIFLTAQAWELSSHKKQ
			ļĮ	ļ		GPGLLSRMGQTVRAVASSMRGVKNRPEEFM
						EMNNFIELFSQKINLIDKISQRIYKEEREYFDE
		l				MKEYGPIHILWSASEEDLVDTLKDVASCIDRC
					<u> </u>	CKATEKRMSGLSEALLPVVHEYVLYSEMLM
	Ì	l			•	GVMKRRDQIQAELDSKVEVLTYKKADTDLL
		ļ				PEEIGKLEDKVECANNALKADWERWKQNM
·						QNDIKLAFTDMAEENIHYYEQCLATWESFLT
624	1004		4700	401	160	SQTNLHLEEASEDKP
634	1984	A	4708	421	158	SYWVGEDYTYKFFEVILIDPFHKAIRRNPDTQ
	Ì	- 1		1		WISKAVYKHREMCGLTSTGRKSHGLEKDRM
625	1005	<u> </u>	4700	42	241	FPHAIGGSCRAA*RRKTLQFPCYH
635	1985	Α	4709	42	341	YIKQPDAKERRRTVHWKKETESEASEITIPPST
	1					PGVPQAPGHWEDYGRGDNFYLPH*DPGGIVL
1						WNIFNRMPIARKNITDGEHHEYLIEVPRLFHT SED
636	1986	A	4721	2	351	EKPDHFFPEGTSFIHEPRRPN*GDLVHCLGGIS
"	.,,,,		.,21	~		RSTTVTVA*LMQKLNLSMNDAYYIVIMKMSS
<u> </u>		1				

SEQ ID	LOPO ID	1 1/-4	LOPO	I Della L		
NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	noa	in in	beginning nucleotide	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
eotide	seq-		USSN	location	location	F=Phenylalanine, G=Glycine, H=Histidine,
seq-	uence		09/496	correspondi	corresponding to last amino	I=Isoleucine, K=Lysine, L=Leucine,
uence	uchec		914	ng to first	acid residue	M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
denoc			714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		1	Ì	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	Sequence	/=possible nucleotide deletion, \=possible
	1			sequence		nucleotide insertion
	 		 -	sequence	 	ISPNFNSMDQPLDFQRTLGLRSPCYNRVPAQK
		ĺ				MYFTTPSNHNAYQVDSVQST
637	1987	A	4726	664	253	NTGLTCSIQRKCGETQLYRREENRLILLLODH
057	1507	^	17/20	004	233	LKSESFQVLTLSPRLEFSGLISAHCNLRLPGSS
				1		DSSASSSRAAGITGVHHHAWLIFFFLVETGFL
	1			1		HAG*AGLELLTSGDPPASASRSAGITGVSHHA
ļ	ľ					RPRETRFL
638	1988	A	4734	24	592	GGMDSRVSGTTSNGETKPVYPVMEKKEEDG
""	1,700	,,	1/31	{ 	1 392	TLERGHWNNKMEFVLSVAGEIIGLGNVWRFP
	İ	1	1			YLCYKNGGGAFFIPYLVFLFTCGIPVFLLETAL
	1	į				GQYTSQGGVTAWRKICPIFEGIGYASOMIVIL
	1					LNVYYIIVLAWALFYLFSSFTIDLPWGGCYHE
j	}	ļ	}	ļ	}	WNTEHCMEFQKTNGSLNGTSENATSPVIEFW
639	1989	A	4743	1040	699	QGLTLLPRMECSATITAHCSLELPGSIDLPTSA
""	1505	**	17.13	1040	""	S*VARTTGTHHHPWLILVLLL*TWGSYYVAQ
			ì			AGLELLGSSNLPAAMVSOSAOIIGHDHCAWA
						TSNHVLYTQEGLRRGKEG
640	1990	A	4771	527	2	GRIDCPHPATVLAQPIFIDACSVLGAYQGAQN
	1		'''	327	-	WIRRRPCLPSGCLKMNREIGPLQHSLCCPGWS
	į .			•		QTPGLKAILLRQPPK*LGLQMESHSCPPAWSA
1	1	l	ł	İ	}	MARSRLTATSASQVQAILLPQPPGTTDSCSPS
ĺ]	1			PDHEQQPLSWVLPPPQKDMNPREQQVALGP
i	Ì	1	1			QAAALPWAVWRNDCFPR
641	1991	A	4780	16	473	RPSSQCGGIPTGWKKGLAPELSSELSSPPLPAR
ĺ		1			''"	LQLAASPYFSPSWAECPQPVPAGTHATWCLA
ĺ	1	[!	ĺ	[RVWARMTPPGPAGIPSHPLPPPPPERSVPIPSP
ĺ						FPARDSGSRQGHSTDRYKHTDAPRDAHRRVP
	1		j.			QRDTDTGVHTGSGTHTHAHTPPEK
642	1992	A	4798	1	487	GYSFRCDIVDYSRSPTALRMARTCWLYYFSK
1	 	l			,	FIELLDTIFFVLRKKNSQVTFLHVFHHTIMPW
	l					TWWFGVKFAAGGLGTFHALLNTAVHVVMY
		ł				SYYGLSALGPAYOKYLWWKKYLTSLOLVOF
	1	}				VIVAIHISQFFFMEDCKYQFPVFACIIMSYSFM
l_						FLLLFLH
643	1993	A	4799	2	391	LMAFIEMHISGSLVYLKIKTKIYSYFSMLNFLL
		l				QEIPLSEILRISSPRDFTNISQGSNPHCFEIITDT
		ļ				MVYFVGENNGDSSHNPVLAATGVGLDVAQS
		ì	1		-	WEKAIRQALMPVTPQASVCTSPGQGKDHSK
						Q*ASVCTSPGQGKDHSKQ
644	1994	Α	4800	488	101	AYPLFAVHPVHTECVAGVVGRAYLLCALFFL
I		l				LSFLGYCKAFRESNKEGAHSSTFWVLLSIFLG
Į]	[]			AVAMLCKEQGITVLVRAATWLGPAFSVCPFP
						SYKDIWGWPCLCGVLHAYIPLLV
645 .	1995	Α	4805	458	126	LLWTTVLCQTPARPQSTMIHLGHILFLLLLPV
]					AAAQTTPGERSSLPAFYPGTSGSCSGCGSLSL
	[Í I	ĺ		PLLAGLVAADAVASLLIVGAVFLCARPRRSP
						AQEDGKVYINMPGRG
646	1996	A	4817	47	1033	LQGDTWHLSFLSHFSRLHGGVPGRGLLEGNL
			j			LQPQAPGHDMTSIPFPGDRLLQVDGVILCGLT
	}					HKQAVQCLKGPGQVARLVLERRVPRSTQQC
						PSANDSMGDERTAVSLVTALPGRPSSCVSVT
						DGPKF*SSN*KRIANGLGFSFVQMEKESCSHL.
			!			KSDLVRIKRLFPGHPAEENGAIAAGDIILGRE
			!			WEGPRKASSSRCRGSWAMQLSVQAGPSFAS
					ľ	YYPAAVEVLHLLRGAPQEVTLLLCRPPPGAL
	i		1			PELEQEWQTPELSADKEFTRATCTDSCTSPIL
i i				Į.		
						GSRGQLGGTVPPQMQGKAWGLRPESSQKAIR
647	1997	Ā	4854	1044	335	

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion IVMPTYDLTDSVLETMGRVSLDMMSVQANT GPPWESKNSTAVWRGRDSRKERLELVKLSRK
						HPELIDAAFTNFFFFKHDENLYGPIVKHISFFD FFKHKYQINIDGTVAAYRLPYLLVGDSVVLK QDSIYYEHFYNELQPWKHYIPVKSNLSDLLEK LKWAKDHDEEAKKIAKAGQEFARNNLMGD DIFCYYFQTFPRNMPIYK
648	1998	A	4867	2030	837	AGMLPAVGSADEEEDPAEEDCPELVPMETTQ SEEEEKSGLGAKIPVTIITGYLGAGKTTLLNYI LTEQHSKRVAVILNEFGEGSALEKSLAVSQG GELYEEWLELRNGCLCCSVKDNGLRAIENLM QKKGKFDYILLETTGLADPGAVASMFWVDA ELGSDIYLDGIITIVDSKYGLKHLAEEKPDGLI NEATRQVALADAILINKTDLVPEEDVKKLRT TIRSINGLGQILETQRSRVDLSNVLDLHAFDSL SGISLQKKLQHVPGTQPHLDQSIVTITFDVPG NAKEEHLNMFIQNLLWEKNVRNKDNHCMEV IRLKGLVSIKDKSQQVIVQGVHELYDLEETPV SWKDDTERTNRLVLLGRNLDKDILKQLFIAT VTETEKQWTTHFKEDQVCT
649	1999	A	4873	226	189	DGVSLLIPKLGVQWAQYWAHWQPPLPGFKR FSCLSLRSSWD*KCAPPHPAFVFLVEMGFHRV GQAGLELRTSGDPPASASQSAGITGVSHLA*P TSMPLLPFQRLCVYI
650	2000	A	4874	2	437	FFFLRRSFAFVAQAGVQWCDLGSPQPLPPGF K*FSCLSLPSSWDYRHAPPPCPS*FLYF**RQG FTMLARLVLNS*PHDLPTSPSQSAEIKGVSHR CPASFYLFLKYYLEAKFCA*GECAPSAGVGA GYKRGHKSCLLINCVVQI
651	2001	A	4898	1701	771	DAWGPETRLARILNPDSFIEPRPGRLPELEATR PHMEPKASCPAAAPLMERKFHVLVGVTGSV AALKLPLLVSKLLDIPGLEVAVVTTERAKHFY SPQDIPVTLYSDADEWEMWKSRSDPVLHIDL RRWADLLLVAPLDANTLGKVASGICDNLLTC VMRAWDRSKPLLFCPAMNTAMWEHPITAQQ VDQLKAFGYVEIPCVAKKLVCGDEGLGAMA EVGTIVDKVKEVLFQHSGFQQS*PGISVMGVP LYSEWVQAKSVKMDVGKIGGYPHLLNGGPA LSLPRGQACSRLNWTEGPGLSFFQPGEAAA
652	2002	A	4927	1	611	FRGRQTSRPARGFSPWRPPGTMQEPSSGECPA SP*LPCASNRLAFGGLIFPCAPLVPYPAPFSPLL PAFSCAPRPRAHTHSRTHPSAPLVPKPSSRAR GQSPIPSRASSPSCSWAQVPGVALARCAGVC KPGDSWRVAACISGRCCSRGRRRGSGPRNPE QSFRGAWGPSFWGSWKSQRELSAGGAQAWP LLGSAGSGLRGEA
653	2003	A	4965	2	283	FFFFI*DGVSLCHPGWNAVARSWLTATSASR VQAVSCFRLPSSWDYRHATMPG*FF*YF**R WGFTHLAILVLNS*PQVICPPWPPKVLTLQA
654	2004	A	4968	3 .	437	RPGIPGRRFRRSWFCQLP*EPEPGLESLATPGD IPAVGLGALGVIPPVRVPQRPPTQRSQGRGW DPERDPGCRVQVSRGPRFGEQKTPGLQGCLP PPCLTHLAAASCVVVWCGRWKRDSAECQCD HSCSAVSQQEDRCRSSSCS
655	2005	A	4983	201	397	MNNNTTCIQPSMISSMALPIIYILLCIVGVFGN TLSQWIFLTKIGKKTSTHIYLSHLVTANLLVC
656	2006	A	4988	332	159	LVHKDMYREFFEEEAQASNKHVTRCLTSLVI REVHIKTMR*HFLPIRLEKNKNNIKD
657	2007	В	5008	129	465	MAGMKTASGDYIDSSWELRVFVGEEDPEAES

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide	Predictéd end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion VTLRVTGESHIGGVLLKIVEQINRKQDWSDH AIWWEQKRQWLLQTHWTLDKYGILADARLF FGPQHRPVILRLPNRRALRLX*
658	2008	A	5017	1	292	FFFFKETESHSVTQAGVQWHDLGSLQPPPPGF KRFSCLSLLSSWDYRCAPPHPANFVFLVETGF HHVAQAGLKLLTL*SANLGLSTSLPIPLFILLS
659	2009	A	5018	17	338	RGHGGKSLTGGTPGNWGDGLLVSEDWSHLIF T*NSLVSPVLGKWSPCLQGPGLSAVHTWPWL MAACWAVHVKTHMRPGLAVLPRLVLNSWS *AIILLWPPKALGLQA
660	2010	A	5028	2	310	SRVDDFVGERRGGCDECLCGHRGLRAVPLG HPGHLCLQPPGGPA*FLDYCRGCCPHPVPGST AGSCPRQKKTTPGPTVLCVCSFWIYQRGEPH HRTGARWNH
661	2011	A	5050	752	431	RQSCSSTQAKVQWFHYGPLQSQPPGLKQSSQ LSLPNSRDHRHVPPRLAIFSFAETGSPYFAQAS LELLGSSHPPTSASQSARITGVSHRAWPLK*F NLNQYQTLTMN
662	2012	A	5054	48	103	ELNNGPFQMPLCNGGNLAVTGSWADRSPLH EAASQGRLLALRTLLSQGYNVNAVTLDHVTP LHEACLGDHVACARTLLEAGANVNAITIDGV TPLFNACSQGSPSCAELLLEYGAQAQLESCLP SPTHEGASKGHHECLDILISWGIDVDQEIPHSG TPLYVACMAQQFHCIWNLIYAGAGVRKGKY
663	2013	Α	5066	951	580	WDTPLPGAGHQSTQKLE*LFAMVEIWQ VRNS*SFAHCASVYKHHYMDGQTPCLFVSSK ADLPEGVAVSGPSPAEFCRKHRLPAPVPFSCA GPAEPSTTIFTQLATMAAFPHLVHAELHPSSF WLRGLLGVVGAAVAAVLSFSLYRVLVKSO
664	2014	Α	5071	550	1	LSFIEVLSMEQVNKTVVREFVVLGFSSLARLQ
		,				QLLFVIFLLLYLFTLGTNAIIISTIVLDRALHTP MYFFLAILSCSEICYTFVIVPKMLVDLLSQKK TISFLGCAIQMFSFLFFGSSHSFLLAAMGYDR YMAICNPLRYSVLMGHGVCMGLMAAAWAC GFTVSLVTTSLVFHLPFHSSNQHE
665	2015	A	5074	496	692	QQYHNTGSAGHHAHCQVGHSPHVHYPSGCG PL*IQRGLPSFNSLEGHSLKDSGHEESVQLDSE HDVQRSLYCDTAVNDVLNTSVTSMGSQMPD HDQNEGFHCREECRILGHSDRCWMPRNPMPI RSKSPEHVRNIIALSIEATAADVEAYDDCGPT KRTFATFGKDVSDHPAEERPTLKGKRTVDVT ICSPKVNSVIREAGNGCEAISPVTSPLHLKSSL PTKPSVSYEIVDPGITARRC
666	2016	A	5080	408	248	IMLLSTSS*VYFQSSTKDSHFFLFDFQKTGPPL VGPKAQLSGLQLQPCLYKRR
667	2017	A	5081	129	247	DLTNSHFFLFDFQKTGPPLGGPKAQFSSLQLQ PCVY*RR
668	2018	A	5086	852	233	NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPTTFADIENLKYLLFTRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSE TCADQSVIWKLSEDKQLAICLKYAGVHAENA EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG CCSDMAITFNGLTPQKMEVMMYGLYRLRAF GHYFNDTLVFLPPVGSEND
669	2019	A	5101		329	PGRPTRPPLLTLLAHVSPEPAGPSCDSLAQPG ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPP RGCQHEAAPCPRGPGSDGLHHASAACASLPP SPILPVLLPELGPL
670	2020	Α	5102	3	547	DAWGNRCAVGAAPRLIHLHLCCTPADPSRKP

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide]	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	Ï			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		ŀ		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		}		peptide		/=possible nucleotide deletion, \=possible
<u> </u>	 			sequence		nucleotide insertion DEL*NMNGRVDYLVTEEEINLTRGPSGLGFNI
						VGGTDQQYVSNDSGIYVSRIKENGAAALDGR
		•		:		LQEGDKILSVNGQDLKNLLHQDAVDLFRNA
	Ì			, 		GYAVSLRVQHRLQVQNGPIGHRGEGDPSGIPI
						FMVLVPVFALTMVAAWAFMRYROOL
671	2021	A	5105	672	400	RDGREELCLQQEPTLPSRICSSAPLLYFLFICPF
						VLLLLLISLLCLYWKARKLSTLRSNTRKEKA
-						LWVDLKEAGGVTTNRMED*EEDECN
672	2022	Α	5148	72	314	IIYFSYNIFLKITELLNDVERLKQALNGLSQLT
	ļ	ļ				YTSGNPTKRQSQLIDTLQHQVKSLEQQLAVS
			1			NQAHGALQEYVLAPCS
673	2023	A	5152	210	335	REILCSRIGRLNIV*MSLFPNLTCRLNAIPIKIPA
			L.			NHFVEVT
674	2024	Α	5153	3	2953	LTEDQPFDILQKSLQEANITEQTLAEEAYLDA
						SIGSSQQFAQAQLHPSSSASFTQASNVSNYSG
1			l			QTLQPIGVTHVPVGASFASNTVGVQHGFMQH
						VGISVPSQHLSNSSQISGSGQIQLIGSFGNHPS
						MMTINNLDGSQIILKGSGQQAPSNVSGGLLV
						HRQTPNGNSLFGNSSSSPVAQPVTVPFNSTNF
						QTSLPVHNIIQRGLAPNSNKVPINIQPKPIQM
						GQQNTYNVNNLGIQQHHVQQGISFASASSPQ
			1			GSVVGPHMSVNIVNQQNTRKPVTSQAVSSTG
						GSIVIHSPMGQPHAPQSQFLIPTSLSVSSNSVH
						HVQTINGQLLQTQPSQLISGQVASEHVMLNR NSSNMLRTNQPYTGPMLNNQNTAVHLVSGQ
						TFAASGSPVIANHASPQLVGGQMPLQQASPT
					•	VLHLSPGQSSVSQGRPGFATMPSVTSMSGPSR
						FPAVSSASTAHPSLGSAVQSGSSGSNFTGDQL
1		l	ĺ			TQPNRTPVPVSVSHRLPVSSSKSTSTFSNTPGT
1						GTQQQFFCQAQKKCLNQTSPISAPKTTDGLR
						QAQIPGLLSTTLPGQDSGSKVISASLGTAQPQ
						QEKVVGSSPGHPAVQVESHSGGQKRPAAKQ
1						LTKGAFILQQLQRDQAHTVTPDKSHFRSLSD
						AVQRLLSYHVCQGSMPTEEDLRKVDNEFETV
						ATQLLKRTQAMLNKYRCLLLEDAMRINPPAE
						MVMIDRMFNQEERASLSRDKRLALVDPEGFQ
1 1						ADFCCSFKLDKAAHETQFGRSDQHGSKASSS
						LQPPAKAQGRDRAKTGVTEPMNHDQFHLVP
			l i	İ		NHIVVSAEGNISKKTECLGRALKFDKVGLVQ YQSTSEEKASRREPLKASQCSPGPEGHRKTSS
					-	RSDHGTESKLSSILADSHLEMTCNNSFQDKSL
						RNSPKNEVLHTDIMKGSGEPQPDLQLTKSLET
						TFKNILELKKAGRQPQSDPTVSGSVELDFPNF
						SPMASQENCLEKFIPDHSEGVVETDSILEAAV
				ĺ		NSILEC PRODUCTION OF THE PROPERTY OF THE PROPE
675	2025	Α	5154	599	1880	LKKMEPFSCDTFVALPPATVDNRIIFGKNSDR
						LYDEVQEVVYFPAVVHDNLGERLKCTYIEID
	l					QVPETYAVVLSRPAWLWGAEMGANEHGVCI
	Ì					GNEAVWGREEVCDEEALLGMDLVRLGLERA
						DTAEKALNVIVDLLEKYGQGGNCTEGRMVF
				j	<u> </u>	SYHNSFLIADRNEAWILETAGKYWAAEKVQE
						GVRNISNQLSITTKIAREHPDMRNYAKRKGW
					İ	WDGKKEFDFAAAYSYLDTAKMMTSSGRYCE
					İ	GYKLLNKHKGNITFETMMEILRDKPSGINME
				1	l	GEFLTTASMVFILPQDSSLPCIHFFTGTPDPER
	1			ł	ľ	SVFKPFIFVPHISQLLDTSSPTFELEDLVKKKS
						HFKPDRRHPLYQKHQQALEVVNNNEEKAKI
						MLDNMRKLEKELFREMESILQNKHLDVEKIV
						NLFPQCTKDEIQIYQSNLSVKVSS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
676	2026	A	5155	sequence 2	306	nucleotide insertion FFFLRRSLALSPRPDCGLQWRNLGSLQAPPPG FTPFSCLSLPSSWDYRRPPPRPANFLYF**RRG FTLLARMVSIS*PHDPPASASQSAGITGVSHRA RPT
677	2027	A	5167	97	740	FFHSVDLLALEQSKTFYKPDWFDIVESEVKCC KEAVCVIDMSSFTEFEITSTGDQALEVLQYLF SNDLDVPVGHIVHTGMLNEGGGYENDCSIAR LNKRSFFMISPTDQQVHCWAWLKKHMPKDS NLLLEDVTWKYTALNLIGPRAVDVLSELSYA PMTPDHFPSLFCKEMSVGYANGIRVMSMTHT GEPGFMLYIPIEYRWGFTMLSTLVSNS
678	2028	Α	5183	1919	2018	PALCRLRDDMTVCVADFGLSKKIYSGDYYRQ GRIAKMPVKWIAIESLADRVYTSKSDVWAFG VTMWEIATRGMTPYPGVQNHEMYDYLLHG HRLKQPEDCLDELCKI**SPQSP
679	2029	A	5190	39	499	RESQVKHFKMRKIDLCLSSEGSEVILATSSDE KHPPENIIDGNPETFWTTTGMFPQEFIICFHKH VRIERLVIQSYFVQTLKIEKSTSKEPVDFEQWI EKDLVHTEGQLQNEEIVAHDGSATYLRFIIVS AFDHFASVHSVSAEGTVVSNLSS
680	2030	A	5204	541	92	EILAVLKLACGDISLNALALMVATAVLTLAPL LLICLSYLFILSAILRVPSAAGRCKAFSTCSAH RTVVVVFYGTISFMYFKPKAKDPNVDKTVAL FYGVVTPSLNPIIYSLRNAEVKAAVLTLLRGG LLSRKASHCYCCPLPLSAGIG
681	2031	A	5207	10	247	VPDNGDVTKLPVCSTLVEETSLTVSEAMEQSI KNESPLPGTLAHTCNTSTLGGRGRWIT*GREF DTSMANMVKPCLYRK
682	2032	A	5210	2	231	FFFETESYSITQAGVQWPNLSSLKTLPPGFK*F SCLSLPSSWDYRCLPPCPANFCIFSRNGVLPC WPGWSRTPDLS
683	2033	A	5218	85	402	CPSVSGLIKSDLRRHNINIGITNVDVKAVSNIF MIILLRSMYRINVKPYFFI*LFFSRVNC*SVIIG YARCYTFLIF*LFL*IPADSPTDQEPKTVMLSK OSESAI
684	2034	A	5220	1	194	NLMKEMQNLNSENHKTWEEYKDTK*IMSYF YG*ALNVIKMAVLPKLMYRFSATLVKIPQHL TDS
685	2035	A .	5228	260	440	LHSQDGNSDPRKPQGEMSAHAFPVQTCGEED QKKTPQVPINFTELSKCS*S*KIMSGERE
686	2036	Α	5239	79	508	GGEAAARAAKLSSPRPHRVGRRERGVGGMS AFSEAALEKKLSELSNSQQSVQTLSLWLIHHR KHSRPIVTVWERELRKAKPNRKLTFLYLAND VIQNSKRKGPEFTKDFAPVIVEAFKHVSSETD ESCKKHLGRVLSIWEERS
687	2037	A	5244	1	428	MAAVVAATALKGRGARNARVLRGILAGATA NKASHNRTRALQSHSSPEGKEEPEPLSPELEYI PRKRGKNPMKAVGLAWAIGFPCGILLFILTKR EVDKDRVKQMKARQNMRLSNTGEYESQRFR ASSQSAPSPDVGSGVQT
688	2038	A	5249	1	1407	LQQTEDKSLLNQGSSSEEVAGSSQKMGQPGP SGDSDLATALHRLSLRRQNYLSEKQFFAEEW QRKIQVLADQKEGVSGCVTPTESLASLCTTQS EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRPGVITKGFTQLPGD AIYHISDLEEDEEEGITFQVQQPLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS GSSSSNTAVNSPALAYRLSIGESIINRRDSTTT

SEO ID	SEQ ID	Met	SEQ	Dendinta d	1 p. 414 1 . 1	
NO: of	NO: of	hod	ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	1	in	nucleotide	location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine.
eotide	seq-	1	USSN	location	corresponding	
seq-	uence	1	09/496	correspondi		M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	1	1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	Ì	1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	}	1		peptide	1 .	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
1		1				FSSTMSLAKLLQERGISAKVYHSPISENPLQPL
l		-	1	1		PKSLAIPSTPPNSPSHSPCPSPLPFEPRVHLSEN
	1	İ	1	t		FLASRPAETFLQEMYGLRPSRNPPDVGQLKM
			-		}	NLVDRLKRLGIARVVKNPGAQENGRCQEAEI
			ĺ			GPQKPDSAVYLNSGSSLLGGLRRNQSLPVIM
100		ļ			<u> </u>	GSFAAPVCTSSPKMGVLKED
689	2039	Α	5254	2	2621	LSLFGSRALGRSGARAMAKAKKVGARRKAS
	ļ	1		ļ		GAPAGARGGPAKANSNPFEVKVNROKFOILG
		1				RKTRHDVGLPGVSRARALRKRTOTLLKEYKE
1		l			1	RDKSNVFRDKRFGEYNSNMSPEEKMMKRFA
	Į	l	1		1	LEQQRHHEKKSIYNLNEDEELTHYGOSLADIE
						KHNDIVDSDSDAEDRGTLSGELTAAHFGGGG
.1				1		GLLHKKTQQEGEEREKPKSRKELIEELIAKSK
1		ł	1	ł		QEKRERQAQREDALELTEKLDODWKEIOTI.I.
						SHKTPKSENRDKKEKPKPDAYDMMVRELGF
			į		İ	EMKAQPSNRMKTEAELAKEEQEHLRKLEAE
1			1	ľ	1	RLRRMLGKDEDENVKKPKHMSADDLNDGFV
ł		ł	1			LDKDDRRLLSYKDGKMNVEEDVOEEOSKEA
						SDPESNEEEGDSSGGEDTEESDSPDSHLDLES
						NVESEEENEKPAKEOROTPGKGLISGKERAG
ł						KATRDELPYTFAAPESYEELRSLLLGRSMEEO
						LLVVERIQKCNHPSLAEGNKAKLEKLFGFLLE
				!		YVGDLATDDPPDLTVIDKLVVHLYHLCOMFP
						ESASDAIKFVLRDAMHEMEEMIETKGRAALP
			1			GLDVLIYLKITGLLFPTSDFWHPVVTPALVCL
]]			SQLLTKCPILSLQDVVKGLFVCCLFLEYVALS
						QRFIPELINFLLGILYIATPNKASOGSTLVHPFR
1 1	i		1 1		J	ALGKNSELLVVSAREDVATWOOSSLSLRWA
1			}			SRLRAPTSTEANHIRLSCLAVGLALLKRCVLM
]	,					YGSLPSFHAIMGPLRALLTDHLADCSHPOELO
						ELCOSTLTEMESOKOLCRPLTCEKSKPVPLKI
1	ļ					FTPRLVKVLEFGRKQGSSKEEOERKRLIHKHK
	j] }			REFKGAVREIRKDNOFLARMOLSEIMERDAF
						RKRKVKQLFNSLATQEGEWKALKRKKFKK
690	2040	Α	5261	1	304	FFFFVFLVETGFHHVGQAGLELLTSGDPPTW
1 1	1		1 1			ASQSAGITGVSHCSWPVIYVLSTLLHAVRNVL
]			J I			FKRTFPLKSSSFLSYDKEIFPILIVLKFYLVTLT
						SFVK
691	2041	Α	5270	3	158	NCHTTHCTANWVHLPGTPPGWKIDGPAAAL
	2010					EVLSSFFFFFLKFSYKPQNIV
692	2042	Α	5282	56	1268	GMEPVGCCGECRGSSVDPRSTFVLSNLAEVV
	j					ERVLTFLPAKALLRVACVCRLWRECVRRVLR
ľ	1			l	ŀ	THRSVTWISAGLAEAGHLEGHCLVRVVAEEL
	1			1		ENVRILPHTVLYMADSETFISLEECRGHKRAR
]	J		J	J	KRTSMETALALEKLFPKQCQVLGIVTPGIVVT
						PMGSGSNRPQEIEIGESGFALLFPQIEGIKIOPF
		- 1		ì	1	HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV
		}		ļ	1	FGYNCCKVGASNYLQQVVSTFSDMNIILAGG
						QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI
			1		1	QSATVLLNEDVSDEKTAEAAMQRLKAANIPE
- 1		l	1	}		IINTIGFMFACVGRGFQYYRAKGNVEADAFR
j		J]	J	KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN
						EVKDDDLFHSYTTIMALIHLGSSK
693	2043	A	5301	362	507	EEIKERFGPGLVIYWYGFIQELDCNRERGILLK
(0)						ACFPTNIVTLCHSIA
694	2044	A	5310	1	204	RVLTAINHTLKENLRKFYKGKKDKPLDLRPK
				İ		KTRAMRRRLNMHEENLKTKKQHRKERLYPL
	22.5	لــــــ				RKYAAKA
695	2045	A	5315	125	1596	ETRSTAVKSEVQVCISLLLCLEDRTMPKKAKP

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, I=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion TGSGKEEGPAPCKQMKLEAAGGPSALNFDSP SSLFESLISPIKTETFFKEFWEQKPLLIQRDDPA LATYYGSLFKLTDLKSLCSRGMYYGRDVNV CRCVNGKKKVLNKDGKAHFLQLRKDFDQKR ATIQFHQPQRFKDELWRIQEKLECYFGSLVGS NVYITPAGSQGLPPHYDDVEVFILQLEGEKH WRLYHPTVPLAREYSVEAEERIGRPVHEFML KPGDLLYFPRGTIHQADTPAGLAHSTHVTIST YQNNSWGDFLLDTISGLVFDTAKEDVELRTG IPRQLLLQVESTTVATRRLSGFLRTLADRLEG TKELLSSDMKKDFIMHRLPPYSAGDGAELSTP GGKLPRLDSVVRLQFKDHIVLTVLPDQDQSD ETQEKMVYIYHSLKNSRETHMMGNEEETEFH GLRFPLSHLDALKQIWNSPAISVKDLKLTTDE EKESLVLSLWTECLIQVV
696	2046	A	5318	1476	742	LMKXYLEAAELGEISDIHTKLLRLSSSOGTIET SLQDIDSRLSPGGSLADAWAHQEGTHPKDRN VEKLQVLLNCMTEIYYQFKKDKAERRLAYN EEQIHKFDKQKLYYHATKAMTHFTDECVKK YEAFLNKSEEWIRKMLHLRKQLLSLTNQCFDI EEEVSKYQEYTNELQETLPQKMFTASSGIKHT MTPIYPSSNTLVEMTLGMKKLKEEMEGVVKE LAENNHILESGGSLTMDGGLRNVDCL
697	2047	A	5320	244	478	LDYNFFLFEMTFGLVSQAGVQWHDLGSLQPP PPGFKQFSCLSLPSSWDYRHLPPHLANFSREG VSPSWPGWSRTPDFR
698	2048	A	5324	266	714	LPIRKSLRSVRSGFPTSQSPITRNLDGTASGSC LAKTVTGSLFRINVGLRGLVAGGIIGALLGTP VGGLLMAFQKYSGETVQERKQKDRKALHEL KLEEWKGRLQVTEHLPEKIESSLQEDEPENDA KKIEALLNLPRNPSVIDKQDKD
699	2049	A	5334	699	277	RPHGHLVCISSSAGLSGVNGLADYCASKFAA FGFAESVFVETFVQKQKGIKTTIVCPFFIKTGM FEGCTTGCPSLLPILEPKYAVEKIVEAILQEKM YLYMPKLLYFMMFLKSFLPLKTGLLIADYLGI LHAMDGFADQKK
700	2050		5344		614	PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS QTSFRAEYKASGGPSVFQKPVRFQVDISSSEG PEPSPRRDGSGGGGIYSVTFTLISGPSRRFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSSPRRGPPKDKK LLATNGTPL
701	2051	A	5346	3	1383	HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTILE GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV IADVKLVADFQRYILYWRKRFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL
702	2052	A	5356	2502	1540	MAAATRGCRPWGSLLGLLGLVSAAAAAWD

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valline, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LASLRCTLGAFCECDFRPDLPGLECDLAQHL AGQHLAKALVVKALKAFVRDPAPTKPLVLSL HGWTGTGKSYVSSLLAHYLFQGGLRSPRVH HFSPVLHFPHPSHIERYKKDLKSWVQGNLTA CGRSLFLFDEMDKMPPGLMEVLRPFLGSSWV VYGTNYRKAIFIFISNTGGEQINQVALEAWRS RRDREEILLQELEPVISRAVLDNPHHGFSNSGI MEERLLDAVVPFLPLQRHHVRHCVLNELAQL GLEPRDEVVQAVLDSTTFFPEDEQLFSSNGCK TVASRIAFFL
703	2053	Α	5380	278	657	LFLQKLRMKTEEEARTHTEIEMFLRKEQQKL EERLEFWMEKYDKDTEMKQNELNALKATKA SDLAHLQDLAKMIREYEQVIIEDRIEKERSKK KVKQDLLELKSVIKLQAWWRGTMIRREIGGF KM
704	2054	A	5381	1	1003	FRGRAVKMAAVVEVEVGGGAAGERELDEV DMSDLSPEEQWRVEHARMHAKHRGHEAMH AEMVLILIATLVVAQLLLVQWKQRHPRSYN MYTLFQMWVVPLYFTVKLHWWRFLVIWILF SAVTAFVTFRATRKPLVQTTPRLVYKWFLLIY KISYATGIVGYMAVMFTLFGLNLLFKIKPEDA MDFGISLLFYGLYYGVLERDFAEMCADYMA STIGFYSESGMPTKHLSDSVCAVCGQQIFVDV SEEGIIENTYRLSCNHVFHEFCIRGWCIVGKK QTCPYCKEKVDLKRMFSNPWERPHVMYGQL LDWLRYLVAWQPVIIGVVQGINYILGLE
705	2055	A	5396	3	675	IYDRDPLQLATRAGQPLDINMAGEPKPYRPKP GNKRPLSALYRLESKEPFLSVGGYVFDYDYY RDDFYNRLFDYHGRVPPPPRAVIPLKRPRVA VTTTRGKGVFSMKGGSRSTASGSTGSKLKS DELQTIKKELTQIKTKIDSVLGRLDKIEKQQK AEAEAQKKLLEESLVLIQEECVSEIADHSTEEP AEGGPDADGEEMTDGIEEAFDEDGGHELFLQ IK
706	2056	A	5410	2	98	GRVGLNLEGRGCSEPKWRHCTPTWATEQDSI
707	2057	Α	5415	6	287	S PFKLTPSFLSHAFSSGQERKVFIELNHIKKCNT VRGVFVLEEFGNYTILLLGLDSHGSNSNLGAP EEGLGAGRKRTSVEKSGGAGVTRKKRDP
708	2058	A	5423	3	291	SSSNPLGSPSTLWKLCSFVLHNKSCCCSFFGS TPTLRAITLTVRVCGFIPEVSKTTNPLGRTNNS GCTIFKTVTLTARSTASLLKSVRPRTHQKE
709	2059	A	5424	679	347	RIRHFEKRGSRGRGRRTSEEDTPKKKKHKGG SEFTDTILSVHPSDVLDMPVDPNEPTYCLCHQ VSYGEMIGCDNPDCPIEWFHFACVDLTTKPK GKWFCPRCVQEKRKKK
710	2060	A	5442	1073	559	QESLKKKIQPKLSLTLSSSVSRGNVSTPPRHSS GSLTPPVTPPITPSSSFRSSTPTGSEYDEEEVDY EESDSDESWTTESAISSEAILSSMCMNGGEEK PFACPVPGCKKRYKNVNGIKYHAKNGHRTQI RVRKPFKCRCGKSYKTAQGLRHHTINFHPPV SAEIIRKMQQ
711	2061	A	5449	1	319	GDSLCVPQYNKYREERVILFLKMASGHAFQP DLVKRIRDAIRMGLSARHVPSLILETKGIPYTL NGKKVEVAVKQIIAGKAVEQGGAFSNPETLD LYRDIPELQGF
712	2062	A	5499	91	749 ·	RPTPGHGDFWMQPLTKDAGMSLSSVTLASAL QVRGEALSEEEIWSLLFLAAEQLLEDLRNDSS DYVVCPWSALLSAAGSLSFQGRVSHIEAAPF

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						KAPELLQGQSEDEQPDASQMHVYSLGMTLY WSAGFHVPPHQPLQLCEPLHSILLTMCEDQPH RRCTLQSVLEACRVHEKEVSVYPAPAGLHIR RLVGLVLGTISEVSREPCFSSSSCWSCVAIKI
713	2063	A	5506	22	478	VEELILVSRLDPHLHTPMYFFLAHLSFLDLSFT TSSIPQLLYNLNGCDKTISYMGCAIQLFLFLGL GGVECLLLAVMAYDRCVAICKPLHYMVIMN PRLCRGLVSVTWGCGVANSLAMSPVTLRLPR CGHHEVDHFLCEMPALIRMACISTV
714	2064	A	5514	25	220	AIRPYWCENNIIGIGKLSTADGKAFADPEVLR RLTSSVSCALDEAAAALTRMRAESTANAGQS DK
715	2065	A	5526	3	810	KVTAPRRPQRYSSGHGSDNSSVLSGELPPAM GRTALFHHSGGSSGYESLRRDSEATGSASSAP DSMSESGAASPGARTRSLKSPKKRATGLQRR RLIPAPLPDTTALGRKPSLPGQWVDLPPPLAG SLKEPFEIKVYEIDDVERLQRPRPTPREAPTQG LACVSTRLRLAERRQQRLREVQAKHKHLCEE LAETQGRLMLEPGRWLEQFEVDPELEPESAE YLAALERATAALEQCVNLCKAHVMMVTCFD ISVAASAAIPGPQEVDV
716	2066	A	5529	458	790	SPGYGENKFTVTSXNIAVPLCEMNKIYSYYSD SSSSERTMDLVLEMCNTNSIHWCGISGRQLG KLHPSSSLCLALTLLSSVQGLQSISGLRLTDTF LKRTYEYDDIAQVCV
717	2067	A	5531	3	460	NSEDLLKYFNPESWQEDLDNMYLDTPRYRG RSYHDRKSKVDLDRLNDDAKRYSCTPRNYS VNIREELKLANVVFFPRCLLVQRCGGNCGCG TVNWRSCTCNSGKTVKKYHEVLQFEPGHIKR RGRAKTMALVDIQLDHHERCDCICSSRPPR
718	2068	A	5586	311	88	AVLKNMAPMTALGLLDLHILNLILFLSAGEDF TSVVSEIMMYILLVFLTLWLLIEMIYCYRKVS KAEEAAQENA
719	2069	A	5598	1	330	KNCANEAVVQKILDRVLSRYDVRLRPNFGSM LATNSTRGLNEDELMAHGQEKDSSSESEDSC PPSPGCSFTEGFSFDLLNPDYVPKVDKWSRFL FPLAFGLFNIVAAERC
720	2070		5628	798	148	LPPAQIPEAWLLLANVVVLILVPLKDRLIDP LLLRCKLLPSALQKMALGMFFGFTSVIVAGV LEMERLHYIHHNETVSQQIGEVLYNAAPLSIW WQIPQYLLIGISEIFASIPGLEFAYSEAPRSMQG AIMGIFFCLSGVGSLLGSSLVALLSLPGGWLH CPKDFGNINNCRMDLYFFLLAGIQAVTALLF VWIAGRYERASQGPASHSRFSRDRG
721	2071	A	5632	146	536	MSALIVRKLRSAELTLFSELPTVLGANVNAA KLHE'I'ALHHAAKVKNVDLIEMLIEFGGNIYA RDNRGKKPSDYTWSSSAPAKCFEYYEKTPLT LSQLCRVNLRKATGVRGLEKIAKLNIPPRLID YLSYN
722	2072		5638	3	3806	CPSLDIRSEVAELRQLENCSVVEGHLQILLMF TATGEDFRGLSFPRLTQVTDYLLLFRVYGLES LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GLLQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT RNSSSIFCHKCEGLCPKECKVGTKTIDSIQAA

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO: in	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide seq- uence	seq- uence		USSN 09/496 914	location correspondi ng to first	corresponding to last amino acid residue	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
				amino acid residue of peptide	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion QDLVGCTHVEGSLILNLRQGYNLEPQLQHSL GLVETITGFLKIKHSFALVSLGFFKNLKLIRGD
		1				AMVDGNYTLYVLDNQNLQQLGSWVAAGLTI PVGKIYFAFNPRLCLEHIYRLEEVTGTRGRQN KAEINPRTNGDRAACQTRTLRFVSNVTEADRI
						LLRWERYEPLEARDLLSFIVYYKESPFQNATE HVGPDACGTQSWNLLDVELPLSRTQEPGVTL ASLKPWTQYAVFVRAITLITTEEDSPHQGAQS
						PIVYLRTLPAAPTVPQDVISTSNSSSHLLVRW KPPTQRNGNLTYYLVLWQRLAEDGDLYLND YCHRGLRLPTSNNDPRFDGEDGDPEAEMESD CCPCQHPPPGQVLPPLEAQEASFQKKFENFLH
						NAITIPISPWKVTSINKSPQRDSGRHRRAAGPL RLGGNSSDFEQEDKVPRERAVLSGLRHFTEY RIDIHACNHAAHTVGCSAATFVFARTMPHRE
						ADGIPGKVAWEASSKNSVLLRWLEPPDPNGL ILKYEIKYRRLGEEATVLCVSRLRYAKFGGV HLALLPPGNYSARVRATSLAGNGSWTDSVAF
						YILGPEEEDAGGLHVLLTATPVGLTLLIVLAA LGFFYGKKRNRTLYASVNPEYFSASDMYVPD EWEVPREQISIIRELGQGSFGMVYEGLARGLE
						AGEESTPVALKTVNELASPRECIEFLKEASVM KAFKCHHVVRLLGVVSQGQPTLVIMELMTR GDLKSHLRSLRPEAENNPGLPQPALGEMIQM AGELADCMAVI AANJENJURDI AARNOMYSO
						AGEIADGMAYLAANKFVHRDLAARNCMVSQ DFTVKIGDFGMTRDVYETDYYRKGGKGLLP VRWMAPESLKDGIFTTHSDVWSFGVVLWEIV TLAEQPYQGLSNEQVLKFVMDGGVLEELEGC
						PLQLQELMSRCWQPNPRLRPSFTHILDSIQEEL RPSFRLLSFYYSPECRGARGSLPTTDAEPDSSP TPRDCSPQNGGPGH
723	2073	A	5672	1	216	LAWIDNILPEKEKKETDKKRKKKGAHEDCD EEPQFPPPSVIKIPMESVQSDPQNGIHCIARKR SSSWSYSL
724	2074	A	5704	4235	940	ARGRRSRPVWAASWGGRGRPAARRRPRGLA ATMGFELDRFDGDVDPDLKCALCHKVLEDP LTTPCGHVFCAGCVLPWVVQEGSCPARCRGR
	,					LSAKELNHVLPLKRLILKLDIKCAYATRGCGR VVKLQQLPEHLERCDFAPARCRHAGCGQVLL RRDVEAHMRDACDARPVGRCQEGCGLPLTH GEORAGGHCCARALRAHNGALQARLGALHK
						ALKKEALRAGKREKSLVAQLAAAQLELQMT ALRYQKKFTEYSARLDSLSRCVAAPPGGKGE ETKSLTLVLHRDSGSLGFNIIGGRPSVDNHDG
						SSSEGIFVSKIVDSGPAAKEGGLQIHDRIIEVN GRDLSRATHDQAVEAFKTAKEPIVVQVLRRT PRTKMFTPPSESQLVDTGTQTDITFEHIMALT
						KMSSPSPPVLDPYLLPEEHPSAHEYYDPNDYI GDIHQEMDREELELEEVDLYRMNSQDKLGLT VCYRTDDEDDIGIYISEIDPNSIAAKDGRIREG DRIIQINGIEVQNREEAVALLTSEENKNFSLLI
						ARAELQLDEGWMDDDRNDFLDDLHMDMLE EQHHQAMQFTASVLQQKKHDEDGGTTDTAT ILSNQHEKDSGVGRTDESTRNDESSEQENNG
						DDATASSNPLAGQRKLTCSQDTLGSGDLPFS NESFISADCTDADYLGIPVDECERFRELLELK CQVKSATPYGLYYPSGPLDAGKSDPESVDKE
						LELLNEELRSIELECLSIVRAHKMQQLKEQYR ESWMLHNSGFRNYNTSIDVRRHELSDITELPE KSDKDSSSAYNTGESCRSTPLTLEISPDNSLRR

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
ŀ				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	ĺ		ĺ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		1		peptide	Soquence	/=possible nucleotide deletion, \=possible
		l		sequence		nucleotide insertion
		 	 	bequeitee		AAEGISCPSSEGAVGTTEAYGPASKNLLSITE
	1		ŀ			DPEVGTPTYSPSLKELDPNQPLESKERRASDG
	1	i				
			į			SRSPTPSQKLGSAYLPSYHHSPYKHAHIPAHA
			[QHYQSYMQLIQQKSAVEYAQSQMSLVSMCK
}]	J				DLSSPTPSEPRMEWKVKIRSDGTRYITKRPVR
İ						DRLLRERALKIREERSGMTTDDDAVSEMKM
ł						GRYWSKEERKQHLVKAKEQRRRREFMMQSR
1						LDCLKEQQAADDRKEMNILELSHKKMMKKR
						NKKIFDNWMTIQELLTHGTKSPDGTRVYNSF
		<u></u>	<u> </u>			LSVTTV
725	2075	Α	5707	3	1770	QISTEVSEAPVANDKPKTLVVKVQKKAADLP
						DRDTWKGRFDFLMSCVGYAIGLGNVWRFPY
1	 	1				LCGKNGGGAFLIPYFLTLIFAGVPLFLLECSLG
1		}	1			QYTSIGGLGVWKLAPMFKGVGLAAAVLSFW
		}			,	LNIYYIVIISWAIYYLYNSFTTTLPWKQCDNP
		1				WNTDRCFSNYSMVNTTNMTSAVVEFWERN
1		1				MHOMTDGLDKPGOIRWPLAITLAIAWILVYF
						CIWKGVGWTGKVVYFSATYPYIMLIILFFRGV
						TLPGAKEGILFYITPNFRKLSDSEVWLDAATO
						IFFSYGLGLGSLIALGSYNSFHNNVYRDSIIVC
ł		ŀ				CINSCTSMFAGFVIFSIVGFMAHVTKRSIADV
i						AASGPGLAFLAYPEAVTQLPISPLWAILFFSM
1.						LLMLGIDSQFCTVEGFITALVDEYPRLLRNRR
						ELFIAAVCIISYLIGLSNITQGGIYVFKLFDYYS
1						ASGMSLLFLVFFECVSISWFYGVNRFYDNIQE
						MVGSRPCIWWKLCWSFFTPIIVAGVFIFSAVQ
						MTPLTMGNYVFPKWGQGVGWLMALSSMVL
						IPGYMAYMFLTLKGSLKQRIQVMVQPSEDIV
L						RPENGPEQPQAGSSTSKEAYI
726	2076	A	5711	156	423	PRRDPGRTPELRGSAPRKTGANMPVRRGHVA
	ļ		<u> </u>			PQNTFLGTIIRKFEGQNKKFIIANARVQNCAII
			L			YCNDGFCEMTGFSRPDVMQKPCTCD
727	2077	Α	5716	3	274	HASEYFFKLCSFQVFLSFPLATIVIDVGLVVIP
						LVKSPNVHYVYVLLLVLSGLLFYIPLIHFKIRL
						AWFEKMTCYLQLLFNICLPDVSEE
728	2078	Α	5737	1899	649	IQASRASPYPRVKVDFALSCHEDLLAPISEPIE
			1			WKYHSPEEEISLGPACWLWDFLRRSQQAGFL
]			LPLSGGVDSAATACLIYSMCCOVCEAVRSGN
			1			EEVLADVRTIVNQISYTPQDPRDLCGRILTTC
			J i		•	YMASKNSSQETCTRARELAQQIGSHHISLNID
1						PAVKAVMGIFSLVTGKSPLFAAHGGSSRENL
1					1	ALQNVQARIRMVLAYLFAQLSLWSRGVHGG
1					1	LLVLGSANVDESLLGYLTKYDCSSADINPIGG
1						
1						ISKTDLRAFVQFCIQRFQLPALQSILLAPATAE LEPLADGOVSOTDEEDMGMTYAELSVYGKL
						RKVAKMGPYSMFCKLLGMWRHICTPRQVAD
1					1	KVKRFFSKYSMNRHKMTTLTPAYHAENYSPE
					i	DNRFDLRPFLYNTSWPWQFRCIENQVLQLER
	0050					AEPQSLDGVD
729	2079	Α	5741	1	5976	PGCAARLSRARAPGPGAAGAGRKRLADPGPP
1				•		PASRRLRAPGSRPRLAPCTRRAAQPAHARMA
				İ		PRAAGGAPLSARAAAASPPPFQTPPRCPVPLL
]	l		LLLLLGAARAGALEIQRRFPSPTPTNNFALDG
				ſ	1	AAGTVYLAAVNRLYQLSGANLSLEAEAAVG
					i	PVPDSPLCHAPQLPQASCEHPRRLTDNYNKIL
1						QLDPGQGLVVVCGSIYQGFCQLRRRGNISAV
						AVRFPPAAPPAEPVTVFPSMLNVAANHPNAS
						TVGLVLPPAAGAGGSRLLVGATYTGYGSSFF
						PRNRSLEDHRFENTPEIAIRSLDTRGDLAKLFT
			L			THE PERSON AND PROPERTY OF THE PERSON OF THE

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Trytophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion FDLNPSDDNILKIKQGAKEQHKLGFVSAFLHP SDPPPGAQSYAYLALNSEARAGDKESQARSL LARICLPHGAGGDAKKLTESYIQLGLQCAGG AGRGDLYSRLVSVFPARERLFAVFERPQGSPA ARAAPAALCAFRFADVRAAIRAARTACFVEP APDVVAVLDSVVQGTGPACERKLNIQLQPEQ LDCGAAHLQHPLSILQPLKATPVFRAPGLTSV AVASVNNYTAVFLGTVNGRLLKINLNESMQ VVSRRVVTVAYGEPVHHVMQFDPADSGYLY LMTSHQMARVKVAACNVHSTCGDCVGAAD AYCGWCALETRCTLQQDCTNSSQQHFWTSA SEGPSRCPAMTVLPSEIDVRQEYPGMILQISGS LPSLSGMEMACDYGNNIRTVARVPGPAFGHQ IAYCNLPRDQFPPFPPNQDHVTVEMSVRVN GRNIVKANFTIYDCSRTAQVYPHTACTSCLSA QWPCFWCSQQHSCVSNQSRCEASPNPTSPQD CPRTLLSPLAPVPTGGSQNILVPLANTAFFQG AALECSFGLEEIFEAVWNESVRCDQVVLH TTRKSQVFPLSLQLKGRPARFLDSPEPMTVM VYNCAMGSPDCSQCLGREDLGHLCMWSDGC RLRGPLOPMAGTCPAPEIRAIEPLSGPLDGGT
						LLTIRGRNLGRRLSDVAHGVWIGGVACEPLP DRYTVSEIVCVTGPAPGPLSGVVTVNASKE GKSRDRFSYVLPLVHSLEPTMGPKAGGTRITI HGNDLHVGSELQVLVNDTDPCTELMRTDTSI ACTMPEGALPAPVPVCVRFERRGCVHGNLTF WYMQNPVITAISPRRSPVSGGRTITVAGERFH MVQNVSMAVHHIGREPTLCKVLNSTLITCPSP GALSNASAPVDFFINGRAYADEVAVAEELLD PEEAQRGSRFRLDYLPNPQFSTAKREKWIKH HPGEPLTLVHIVSTKGAGKEQDSLGLQSHEY RVKIGQVSCDIQIVSDRIHCSVNESLGAAVGQ LPITIQVGNFNQTIATLQLGGSETAIIVSIVICSV LLLLSVVALFVFCTKSRRAERYWQKTLLQME EMESQIREEIRKGFAELQTDMTDLTKELNRSQ GIPFLEYKHFVTRTFFPKCSSLYEERYVLPSQT LNSQGSSQAQETHPLLGEWKIPESCRPNMEE GISLFSSLLDNKHFLIVFVHALEQQKDFAVRD RCSLASLLTIALHGKLEYYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEPFFLLLCAIKQQINKGSIDAITG KARYTLNEEWLLRENIEAKPRNLNVSFQGCG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEKYFHLVLPTDELAEPKKSH
730	2080	A	5744	3	292	RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KAILSIREDKPPLAVKYFFDFLEEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA QPSPLFHSHLETLQLLRTAQLPEQVSWPWGO
		A			382	VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC FRVDEVNWTTWNTNVGIINEDPGNCEGVKRT

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LSFSLRSSRVSGRHWKNFALVPLLREASARD RQSAQPEEVYLRQFSGSLKPEDAEVFKSPAAS
732	2082	A	5753	198	3	GEK AQAESSTVASPEATAGPLCTRIPNVPPPTPIRP
733	2083	A	5754	2	2223	PGKLQAQLPCPSPVRFTSARIPPASRPQTKS AAGPPGLEAEGRAPESAGPGPGGDAAETPGL PPAHSGTLMMAFRDVTVQIANQNISVSSSTAL SVANCLGAQTVQAPAEPAAGKAEQGETSGR EAPEAPAVGREDASAEDSCAEAGASGAADG ATAPKTEEEEEEETAEVGRGAEAEAGDLEQ LNRTSTSTKSAKSGSEASASASKDALQAMILS LPRYHCENPASCKSPTLSTDTLRKRLYRIGLN LFNINPDKGIQFLISRGFIPDTPIGVAHFLLQRK GLSRQMIGEFLGNSKKQFNRDVLDCVVDEM DFSSMELDEALRKFQAHIRVQGEAQKVERLIE AFSQRYCMCNPEVVQQFHNPDTIFILAFAIIL NTDMYSPNIKPDRKMMLEDFIRNLRGVDDG ADIPRELVVGIYERIQQKELKSNEDHVTYVTK VEKSIVGMKTVLSVPHRRLVCCSRLFEVTDV NKLQKQAAHQREVFLFNDLLVILKLCPKKKS SSTYTFCKSVGLLGMQFQLFENEYYSHGITLV TPLSGSEKKQVLHFCALGSDEMQKFVEDLKE SIAEVTELEQIRIEWELEKQQGTKTLSFKPCGA QGDPQSKQGSPTAKREAALRERPAESTVEVSI HNRLQTSQHNSGLGAERGAPVPPPDLQPSPPR QQTPPLPPPPTPPGTLVQCQQIVKVIVLDKPC LARMEPLLSQALSCYTSSSSDSCGSTPLGGPG SPVKVTHQPPLPPPPPPPNHPHPQFCPPGSLLH
734	2084	A	5788	8	362	GHRYSSGSRSLV SSVMGDLVGQGLEEQIVARDENSWLIDGGTP IDDVMRVLDIDEFPQSGNYETIGGFMMFMLR KIPKRTDSVKFAGYKFEVVDIDNYRIDQLLVT
						RIDSKATALSPKLPDAKDKEESVA
735	2085	A	5827		1257	MVFSAVLTAFHTGTSNTTFVVYENTYMNITL PPPFQHPDLSPLLRYSFETMAPTGLSSLTVNST AVPTTPAAFKSLNLPLQITLSAIMIFILFVSFLG NLVVCLMVYQKAAMRSAINILLASLAFADM LLAVLNMPFALVTILTTRWIFGKFFCRVSAMF FWLFVIEGVAILLIISIDRFLIIVQRQDKLNPYR AKVLIAVSWATSFCVAFPLAVGNPDLQIPSRA PQCVFGYTTNPGYQAYVILISLISFFIPFLVILY SFMGILNTLRHNALRIHSYPEGICLSQASKLGL MGLQRPFQMSIDMGFKTRAFTTILIILFAVFIVC WAPFTTYSLVATFSKHFYYQHNFFEISTWLL WLCYLKSALNPLIYYWRIKKFHDACLDMMP KSFKFLPQLPGHTKRRIRPSAVYVCGEHRTVV
736	2086	A	5870	3	268	FTRSDELARHYRTHTGEKRFSCPLCPKOFSRS DHLTKHARRHPTYHPDMIEYRGRRRTPRIDPP LTSEVESSASGSGPGPAPSFTTCL
737	2087	٨	5871	2	521	LTWPQLFLETLPELLHMSRPAEDGPSPGALVR RSSSLGYISKAEEYFLLKSRSDLMFEKQSERH GLARRLTTARRPPASSEQAQQELFNELKPAV DGANFIVNHMRDQNNYNEEKDSWNRVART VDRLCLFVVTPVMVVGTAWIFLQGVYNQPPP QPFPGDPYSYNVQDKRFI
738	2088	A	5881	1	1160	LVVTAITAILAFPNEYTRMSTSELISELFNDCG LLDSSKLCDYENRFNTSKGGELPDRPAGVGV YSAMWQLALTLILKIVITIFTFGMKIPSGLFIPS MAVGAIAGRLLGVGMEQLAYYHQEWTVFNS

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	nou	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide				location		
	seq-		USSN		corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	[09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		i	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
Ì	1			peptide		/=possible nucleotide deletion, \=possible
	.	↓	<u> </u>	sequence		nucleotide insertion
				!		WCSQGADCITPGLYAMVGAAACLGGVTRMT
	1]		İ		VSLVVIMFELTGGLEYIVPLMAAAMTSKWVA
ł		1		İ		DALGREGIYDAHIRLNGYPFLEAKEEFAHKTL
		1		}		AMDVMKPRRNDPLLTVLTQDSMTVEDVETII
		ļ				SETTYSGFPVVVSRESQRLVGFVLRRDLIISIE
ļ	1	j			,	NARKKQDGVVSTSIIYFTEHSPPLPPYTPPTLK
		ļ.				LRNILDLSPFTVTDLTPMEIVVDIFRKLGLRQC
l	1	ľ			'	LVTHNGRLLGIITKKDVLKHIAQMANQDPDSI
	↓		<u></u>			LFN
739	2089	Α	5892	2	916	TLQLAASVPFFAISLISWWLPESARWLIINGKP
}	1	ļ			}	DQALQELRKVARINGHKEAKNLTIEVLMSSV
					ĺ	KEEVASAKEPRSVLDLFCVPVLRWRSCAMLV
	ł					VNFSLLISYYGLVFDLQSLGRDIFLLQALFGA
	1	1			İ	VDFLGRATTALLLSFLGRRTIQAGSQAMAGL
ĺ		ĺ		1		AILANMLVPQDLQTLRVVFAVLGKGCFGISL
	1					TCLTIYKAELFPTPVRMTADGILHTVGRLGA
1						MMGPLILMSRQALPLLPPLLYGVISIASSLVVL
	1	1	,			FFLPETQGLPLPDTIQDLESQKSTAAQGNRQE
	İ	1				AFTVESTSLLEIVALHGAL
740	2090	A	5900	2	426	RPIKTLGIGFHFSVDGVHFLTQREVQNLWKE
	!			ĺ		NLIILDTAKKHGYEVVDTFTITMGRYKEFLQG
						KCGCHFHEVVKSKLSKEYNFIKMKRSRNHIM
						GRYFSNQSKLQQGTVTNFRSPYHVRGPINQV
						CSEILLSRMCANKRTM
741	2091	Α	5910	3	412	RMPESTLLIICENGYILEAPLPTIKQEEDDHDV
						VSYEIKDMCIKCFHFSSVKSKILRLIEIEKRER
						QRELKEKIREERRNKLAAEMGEDGEKEFQEE
		ł				EEEKEEEEEEEPLPEIFIPSTPSPILCGFYSEPG
		!				KFWV
742	2092	Α	5936	1	482	MGCRLLCCVVFCLLQAGPLDTAVSQTPKYLV
		ŀ				TQMGNDKSIKCEQNLGHDTMYWYKQDSKK
						FLKIMFSYNNKELIINETVPNRFSPKSPDKAHL
						NLHINSLELGDSAVYFCASSODTALOSHCIPV
		-		!		HKPPGSARKLQGSVCTCTQGSSLHSLMASDG
	İ	ĺ	1	' i		VPVC
743	2093	Α	5938	1	1566	MNSFFGTPAASWCLLESDVSSAPDKEAGRER
						RALSVQQRGGPAWSGSLEWSRQSAGDRRRL
						GLSRQTAKSSWSRSRDRTCCCRRAWWILVPA
					İ	ADRARRERFIMNEKWDTNSSENWHPIWNVN
				1		DTKHHLYSDINITYVNYYLHQPQVAAIFIISYF
				ļ		LIFFLCMMGNTVVCFIVMRNKHMHTVTNLFI
						LNLAISDLLVGIFCMPITLLDNIIAGWPFGNTM
				ł	ŀ	CKISGLVQGISVAASVFTLVAIAVDRFQCVVY
					[PFKPKLTIKTAFVIIMIIWVLAITIMSPSAVMLH
						VQEEKYYRVRLNSQNKTSPVYWCREDWPNQ
	}					EMRKIYTTVLFANIYLAPLSLIVIMYGRIGISLF
	ļ					RAAVPHTGRKNQEQWHVVSRKKQKIIKMLLI
	1 1					VALLFILSWLPLWTLMMLSDYADLSPNELOII
					İ	NIYIYPFAHWLAFGNSSVNPIIYGFFNENFRRG
						FQEAFQLQLCQKRAKPMEAYALKAKSHVLIN
]					TSNQLVQESTFQNPHGETLLYRKSAEKPQQE
	[[ſ		LVMEELKETTNSSEI
744	2094	-A	5066	140	227	
/44	2094	A	5966	149	327	SHVCVSHYAGSSGCPAGAGAGAVALGISAVA
	i l		5050	413	856	LYDYQGGRLGVARGAWYMEAPDIRQGDM GAPHTDWAWAPTPMSGLGSGRGRQGTLASS
745	2005	A 1			A 30	CAPHILIWAWAPIPWSCH GSGRCRDCHI ASS
745	2095	A	5970	413	950	
745	2095	A	3970	413	050	PLSLPLLLAGVTGILATELFDQMARPAACMV
745	2095	A	3970	413	050	PLSLPLLLAGVTGILATELFDQMARPAACMV CGALMWIMLILVGLGFPFIMEALSHFLYVPFL
745	2095	A	5970	413		PLSLPLLLAGVTGILATELFDQMARPAACMV

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location; corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
746	2096	Α	5971	3	1343	AQTARRIIGLELDTEGHRLFVAFSGCIVYLPLS RCARHGACQRSCLASQDPYCGWHSSRGCVDI RGSGGTDVDQAGNQESMEHGDCQDGATGSQ SGPGDSAYGVRRDLPPASASRSVPIPLLLASV AAAFALGASVSGLLVSCACRRAHRRRGKDIE TPGLPRPLSLRSLARLHGGGPEPPPPSKDGDA VQTPQLYTTFLPPPEGVPPPELACLPTPESTPE LPVKHLRAAGDPWEWNQNRNNAKEGPGRSR GGHAAGGPAPRVLVRPPPPGCPGQAVEVTTL EELLRYLHGPQPPRKGAEPPAPLTSRALPPEP APALLGGPSPRPHECASPLRLDVPPEGRCASA PARPALSAPAPRLGVGGGRRLPFSGHRAPPAL LTRVPSGGPSRYSGGPGKHLLYLGRPEGYRG RALKRVDVEKPQLSLKPPLVGPSSRQAVPNG GRFNF
747	2097	A	5998	2	754	DHASLPCSWNHRFDVETRHVFIGDHSGQVTI LKLEQENCTLVTTFRGHTGGVTALCWDPVQ RVLFSGSSDHSVIMWDIGGRKGTAIELQGHN DRVQALSYAQHTRQLISCGGDGGIVVWNMD VERQETPEWLDSDSCQKCDQPFFWNFKQMW DSKKIGLRQHHCRKCGKAVCGKCSSKRSSIPL MGFEFEVRVCDSCHEAITDEERAPTATFHDSK HNIVHVHFDATRGWLLTSGTDKVIKLWDMT PVVS
748	2098	A	6001	2	747	AMVFGGVVPYVPQYRDIRRTQNADGFSTYV CLVLLVANILRILFWFGRRFESPLLWQSAIMIL TMLLMLKLCTEVRVANELNARRRSFTAADS KDEEVKVAPRRSFLDFDPHHFWQWSSFSDYV QCVLAFTGVAGYITYLSIDSALFVETLGFLAV LTEAMLGVPQLYRNHRHQSTEGMSIKMVLM WTSGDAFKTAYFLLKGAPLQFSVCGLLQVLV
749	2099	A	6002	2	447	DLAILGQAYAFARHPQKPAPHAVHPTGTKAL GRPDRSELVRMHILEETFAEPSLQATQMKLK RARLADDLNEKIAQRPGPMELVEKNILPVDSS VKEAIIGVGKEDYPHTQGDFSFDEDSSDALSP DQPASQESQGSAASPSEPKVSESPSPVTTNTP AQFASVSPTVPEFLKTPPTAD
750	2100	A	6004	2	427	LLTQAMLVLPHRPQWFTPGPRLQAQGPCQEG WRWELRLRNYVPEDEDLNKRRVPQAKPDAV QEKVKEQLEAAKPEPVIEEVDLAKLAPRKPD WDLKRDVAKKLEKLLKRTQRAIAELIRERLK GQEDSLDSAVDAATEHKTC
751	2101	A	6007		1280	TDQAKVDNQPEKLVRSAEDVSTVPTQPDNPF SHPDKLKRMSKSVPAFLQDESDDRETDTASE SSYQLSRHKKSPSSLTNLSSSSGMTSLSSVSGS VMSVYSGDFGNLEVKGNIQFAIEYVESLKEL HVFVAQCKDLAAADVKKQRSDPYVKAYLLP DKGKMGKKKTLVVKKTLNPVYNEILRYKIEK QILKTQKLNLSIWHRDTFKRNSFLGEVELDLE TWDWDNKQNKQLRWYPLKRKTAPVALEAE NRGEMKLALQYVPEPVPGKKLPTTGEVHIWV KECLDLPLLRGSHLNSFVKCTILPDTSRKSRQ KTRAVGKTTNPIFNHTMVYDGFRPEDLMEAC VELTVWDHYKLTNQFLGGLRIGFGTGKSYGT EVDWMDSTSEEVALWEKMVNSPNTWIEATL PLRMLLIAKISK
752	2102	Α	6028	108	1283	KEIFSPFELISVKPLCLLLGVTCSQSMAFEELL SQVGGLGRFQMLHLVFILPSLMLLIPHILLENF AAAIPGHRCWVHMLDNNTGSGNETGILSEDA

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LLRISIPLDSNLRPEKCRRFVHPQWQLLHLNG TIHSTSEADTEPCVDGWVYDQSYFPSTIVTKW DLVCDYQSLKSVVQFLLLTGMLVGGIIGGHV SDRFGRRFILRWGLLQLAITDTCAAFAPTFPV YCVLRFLAGFSSMIIISNNSLPITEWIRPNSKAL VVILSSGALNIGQIILGGLAYVFRDWQTLHVV ASVPFFVFFLLSRWLVESARWLIITNKLDEGL KALRKVARTNGIKNAEETLNIEVVRSTMQEE LDAAQTKTTVWDLFRNPSMRKRICILVFLRK
753	2103	Α	6043		1470	KNLKEKA DSFESILRLIFEIHHSGEKGDIVVFLACEQDIEK VCETVYQGSNLNPDLGELVVVPLYPKEKCSL FKPLDETEKRCQVYQRRVVLTTSSGEFLIWSN SVRFVIDVGVERRKVYNPRIRANSLVMQPISQ SQAEIRKQILGSSSSGKFFCLYTEEFASKDMTP LKPAEMQEANLTSMVLFMKRIDIAGLGHCDF MNRPAPESLMQALEDLDYLAALDNDGNLSE FGIIMSEFPLDPQLSKSILASCEFDCVDEVLTIA AMVTAPNCFSHVPHGAEEAALTCWKTFLHPE GDHFTLISIYKAYQDTTLNSSSEYCVEKWCRD YFLNCSALRMADVIRAELLEIIKRIELPYAEPA FGSKENTLNIKKALLSGYFMQIARDVDGSGN YLMLTHKQVAQLHPLSGYSITKKMPEWVLF HKFSISENNYIRITSEISPELFMQLVPQYYFSNL PPSESKDILQQVVDHLSPVSTMNKEQQMCET CPETEQRCTLQ
754	2104	A	6055	2	394	YYALHHWPFPDLLCQTTGAIFQMNMYGSCIF LMLINVDRYAAIVHPLRLRHLRRPRVARLLC LGVWALILVFAVPAARVHRPSRCRYRDLEVR LCFESFSDELWKGRLLPLVLLAEALGFLLPLA AVVYSS
755	2105	A	6059		1795	LGLGSGTLLSVSEYKKKYREHVLQLHARVKE RNARSVKITKRFTKLLIAPESAAPEEALGPAEE PEPGRARRSDTHTFNRLFRRDEEGRRPLTVVL QGPAGIGKTMAAKKILYDWAAGKLYQGQVD FAFFMPCGELLERPGTRSLADLILDQCPDRGA PVPQMLAQPQRLLFILDGADELPALGGPEAAP CTDPFEAASGARVLGGLLSKALLPTALLLVTT RAAAPGRLQGRLCSPQCAEVRGFSDKDKKK YFYKFFRDERRAERAYRFVKENETLFALCFV PFVCWIVCTVLRQQLELGRDLSRTSKTTTSVY LLFITSVLSSAPVADGPRLQGDLRNLCRLARE GVLGRRAQFAEKELEQLELRGSKVQTLFLSK KELPGVLETEVTYQFIDQSFQEFLAALSYLLE DGGVPRTAAGGVGTLLRGDAQPHSHLVLTT RFLFGLLSAERMRDIERHFGCMVSERVKQEA LRWVQGQGGCPGVAPEVTEGAKGLEDTEE PEEEEGEEPNYPLELLYCLYETQEDAFVRQA LCRFPELALQRVRFCRMDVAVLSYCVRCCPA GQALRLISCRLVAAQEKKKKSLGKRLQASLG GG
756	2106	A .	6060	12	436	SGRPTRPAKPTGQGMGRFMLTLVCQGSIMMS ARDLIMNNLTELQPGLFHHLRFLEELRLSGNH LSHIPGQAFSGLYSLKILMLHNNQLGGIPAQA LWELPSLQSLRLDANLISLVPERSFEGLSSLRH LWLDDNALTEIPS
757	2107	Α .	6063	54	419 .	ITPLGLGAADMCAFPWLLLLLLLQEGSQRRL WRWCGSEEVVAVLQESISLPLEIPPDEEVENII WSSHKSLATVVPGKEGHPATIMVTNPHYQG

CEO ID	T OF A TO	136	1050	1 B 11 - 1	<u> </u>	· · · · · · · · · · · · · · · · · · ·
SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	, nou	in NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine.
eotide	seq-	ŀ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	j	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline.
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	İ			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	Ì	1	1	peptide		/=possible nucleotide deletion, \=possible
		İ		sequence		nucleotide insertion
				!		QILTMLLRSLQQPSASWPRDCSSSCSW
758	2108	A	6066	125	438	IGISCPATIFVPMFSHSLIGIGEEYQLPYYNMV
	1	ļ		j	1	PSDPSYEDMREVVCVKRLRPIVSNRWNSDEC
		1				LRAVLKLMSECWAHNPASRLTALRIKKTLAK
		1				MVESQDVKI
759	2109	A	6072	3	650	PGRRFRPAALEERAMEKLREKVPFQNRGKGT
	ŀ					LSSIIPNNSDTRKATETTSLSSKPEYVNPDFRW
•						SKDPSSKSGNLLETSEVGWTSNPEELDPIRLA
						LLGKSGLSCQVGSATSHPVSCQEPIDEDQRISP
						KDKSTAGREFSGQVSHQTTSENQCTPIPSSTV
		ĺ	ì	i	ł	HSSVADMQNMPAAVHALLTQPSLSAAPFAQ
760	2110	A	6077	3	730	RYLGTLPSTGSTTLPQCHAGNATVW PLRLTLMEEVILLGLKDREGYTSFWNDCISSG
	2110	^	1 00 //	3	730	LRGCMLIELPLRGRLQLEACGMRRKSLLTRK
						VICKSDAPTGDVLLDEALKHVKETQPPETVQ
						NWIELLSGETWNPLKLHYQLRNVRERLAKNL
				Į		VEKGVLTTEKQNFLLFDMTTHPLTNNNIKQR
				Į		LIKKVQEAVLDKWVNDPHRMDRRLLALIYL
			1			AHASDVLENAFAPLLDEQYDLATKRVRQLLD
		Ĺ	<u> </u>	İ	<u></u>	LDPEVECLKANTNEVLWAVVAAFTK
761	2111	Α	6078	833	390	IVSFHLSGFKKFVRPFSFLSVHGLQVDEYHSV
						HQKLSADMADHSNLIRSLLVGAEDARLMRD
		1	1			MKTMKSRYMELYDLNRDLLNGYKIRWNNH
						TELLGNLKAVNQAIQRAGRLRVGKPKNQVIT
762	2112	A	6079	2	2686	ACRDAIRSNNINTLFKIMRVGTASS
702	2112	^	0073		2000	KKAITCGEKEKQDLIKSLAMLKDGFRTDRGS HSDLWSSSSSLESSSFPLPKQYLDVSSQTDISG
						SFGINSNNQLAEKVRLRLRYEEAKRRIANLKI
						QLAKLDSEAWPGVLDSERDRLILINEKEELLK
						EMRFISPRKWTQGEVEQLEMARKRLEKDLQ
			ļ. <u> </u>			AARDTQSKALTERLKLNSKRNQLVRELEEAT
						RQVATLHSQLKSLSSSMQSLSSGSSPGSLTSSR
						GSLVASSLDSSTSASFTDLYYDPFEQLDSELQ
]				SKVEFLLLEGATGFRPSGCITTIHEDEVAKTQ
						KAEGGGRLQALRSLSGTPKSMTSLSPRSSLSS
	1					PSPPCSPLMADPLLAGDAFLNSLEFEDPELSA
						TLCELSLGNSAQERYRLEEPGTEGKQLGQAV
						NTAQGCGLKVACVSAAVSDESVAGDSGVYE
					ľ	ASVQKLGASEAAAFDSDESEAVGATRIQIALK YDEKNKQFAILIIQLSNLSALLQQQDQKVNIR
						VAVLPCSESTTCLFRTRPLDASDTLVFNEVFW
			1			VSMSYPALHQKTLRVDVCTTDRSHLEECLGG
))		ļ	AQISLAEVCRSGERSTRWYNLLSYKYLKKOS
						RELKPVGVMAPASGPASTDAVSALLEOTAVE
			1			LEKRQEGRSSTQTLEDSWRYEETSENEAVAE
						EEEEEVEEEGEEDVFTEKASPDMDGYPALK
]			VDKETNTETPAPSPTVVRPKDRRVGTPSQGPF
			1 1	ľ		LRGSTIIRSKTFSPGPQSQYVCRLNRSDSDSST
						LSKKPPFVRNSLERRSVRMKRPSPPPQPSSVK
						SLRSERLIRTSLDLELDLQATRTWHSQLTQEIS
			[[`	VLKELKEQLEQAKSHGEKELPQWLREDERFR
					1	LLLRMLEKRMDRAEHMGELQTDKMMRAAA
						KDVHRLRGQSCKEPPEVQSFREKMAFFTRPR
762	2112		6000		1650	MNIPALSADDV
763	2113	Α	6082	3	1558	PHPIRFSKLCVSFNNQEYNQFCVIEEASKANE
						VLENLTQGKMCLVPGKTRKLLFKFVAKTED
ł						VGKKIEITSVDLALGNETGRCVVLNWQGGGG
ļ				ļ]	DAASSQEALQAARSFKRRPKLPDNEVHWGSII IQASTMIISRVPNISVHLLHEPPALTNEMYCLV
;	J		i l	j	i	IVI WITH THE VILLE OF THE PROPERTY OF THE PROP

were with that that the

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Trytophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion VTVQSHEKTQIRDVKLTAGLKPGQDANLTQK THVTLHGTELCDESYPALLTDIPVGDLHPGEQ LEKMLYVRCGTVGSRMFLVYVSYLINTTVEE KEIVCKCHKDETVTIETVFPDVAVKFVSTKF EHLERVYADIPFLLMTDLLSASPWALTIVSSE LHLAPSMTTVDQLESQVDNVILQTGESASECF CLQCPSLGNIEGGVATGHYIISWKRTSAMENI PIITTVITLPHVIVENIPLHVNADLPSFGRVRES LPVKYHLQNKTDLVQDVEISVEPSDAFMFSG LKQIRLRILPGTEQEMLYNFYPLMAGYQQLPS
			ļ			LNINLLRFPNFTNQLLRRFIPTSIFVKPQGRLM
764	2114	A	6093	1	1422	DDTSIAAA AAADLANSNAGAAVGRKAGPRSPPSAPAPAP PPPAPAPPTLGNNHQESPGWRCCRPTLRERN ALMFNNELMADVHFVVGPPGATRTVPAHKY VLAVGSSVFYAMFYGDLAEVKSEIHIPDVEPA AFLILLKYMYSDEIDLEADTVLATLYAAKKYI VPALAKACVNFLETSLEAKNACVLLSQSRLF EEPELTQRCWEVIDAQAEMALRSEGFCEIDR QTLEIIVTREALNTKEAVVFEAVLNWAEAEC KRQGLPITPRNKRHVLGRALYLVRIPTMTLEE FANGAAQSDILTLEETHSIFLWYTATNKPRLD FPLTKRKGLAPQRCHRFQSSAYRSNQWRYRG RCDSIQFAVDRRVFIAGLGLYGSSSGKAEYSV KIELKRLGVVLAQNLTKFMSDGSSNTFPVWF EHPVQVEQDTFYTASAVLDGSELSYFGQEGM TEVQCGKVAFQFQCSSDSTNGTGVQGGQIPE LIFYA
765	2115	Α .	6099	1		SGFTHYAIYDFIVKGSCFCNVHADQCIPVHGF RPVKAPGTFHMVHGKCMCKHNTAGSHCQH CAPLYNDRPWEAADGKTGAPNECRTCKCNG HADTCHFDVNVWEASGNRSGGVCDDCQHN TEGQYCQRCKPGFYRDLRRPFSAPDACKPCS CHPVGSAVLPANSVTFCDPSNGDCPCKPGVA GRRCDRCMVGYWGFGDYGCRPCDCAGSCD PITGDCISSHTDIDWYHEVPDFRPVHNKSEPP WEWEDAQGFSALLHSGKCEKEQTLGNAKA FCGMKYSYVLKIKILSAHDKGTHVEVNVKIK KVLKSTKLKIFRGKRTLYPESWTDRGCTCPIL NPGLEYLVAGHEDIRTGKLIVNMKSFVQHWK PSLGRKVMDILKRECK
766	2116	Α	6103	2	384	MTAAATATVLKEGVLEKRSGGLLQLWKRKR CVLTERGLQLFEAKGTGGRPKELSFARIKAVE
						CVESTGRHIYFTLVTEGGGEIDFRCPLEDPGW NAQITLGLVKFKNQQAIQTVRARQSLGTGTL
767	2117	A	6106	1	542	VS SGSSHASDGSGFQELRICSEDQTPLIAGMCSLP MARYYIIKYADQKALYTRDGQLLVGDPVAD NCCAEKICTLPNRGLDRTKVPIFLGIQGGSRC LACVETEEGPSLQLEDVNIEELYKGGEEATRF TFFQSSSGSAFRLEAAAWPGWFLCGPAEPQQ PVQLTKESEPSARTKFYFEQSW
768	2118	A	6109	3	292	FILQAVLQLSSQEARYKAFGTCVSHIGAILAF YTPSVISSVMHRVARCAAPHVHILLANFYLLF
769	2119	A	6110	1	711	PPMVNPIIYGVKTKQIRDSLGSIPEKGCVNRE RHEPSCSNGVASTKSKQNHSKYPAPSSSSSS SSSSSSSSSSVNYSESNSTDSTKSQHHSSTSNQ ETSDSEMEMEAEHYPNGVLGSMSTRIVNGAY KHEDLQTDESSMDDRHPRRQLCGGNQAATE

-- ---- was well, spell stells diels

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	I Amino said some sa (A-Alemino C-Costai
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1104	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	1		USSN	location		
1	seq- uence	l	09/496	correspondi	corresponding to last amino	I=Isoleucine, K=Lysine, L=Leucine,
seq- uence	uence	1	914	ng to first	acid residue	M=Methionine, N=Asparagine, P=Proline,
uence	1		714	amino acid		Q=Glutamine, R=Arginine, S=Serine,
[residue of	of peptide	T=Threonine, V=Valine, W=Tryptophan,
i	ĺ	ĺ			sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	1			peptide	i	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
					į	RIILFGRELQALSEQLGREYGKNLAHTEMLQD
i	1	ĺ		1		AFSLLAYSDPWSCPVGQQLDPIQREPVCAAL
	l	1	1			NSAILESQNLPKQPPLMLALGQASECLRLMA
						RAGLGSCSFARVDDYLH
770	2120	A	6125	2	570	YFGLNLHVQHLGNNVFLLQTLFGAVILLANC
1	•	[1		ĺ	VAPWALKYMNRRASQMLLMFLLAICLLAIIF
		ļ	1			VPQEMQMLREVLATLGLGASALANTLAFAH
1	Į.	ļ				GNEVIPTIIRARAMGINATFANIAGALAPLMM
j	j	}	J]	ILSVYSPPLPWIIYGVFPFISGFAFLLLPETRNK
1	Ι.	İ				PLFDTIQDEKNERKDPREPKQEDPRVEVTQF
771	2121	Α	6126	909	353	RSFVLDTASAICNYNAHYKNHPKYWCRGYF
1						RDYCNIIAFSPNSTNHVALRDTGNQLIVTMSC
1		ļ	1			LTKEDTGWYWCGIORDFARDDMDFTELIVT
						DDKGTLANDFWSGKDLSGNKTRSCKAPKVV
•						RKADRSRTSILIICILITGLGIISVISHLTKRRRS
i I	l	Ì	l		}	QRNRRVGNTLKPFSRVLTPKEMAPTEQM
772	2122	A	6148	7	810	
''-	2122	^	0148	'	010	FVLGILALSHTISPFMNKFFPASFPNRQYQLLF
						TQGSGENKEEIINYEFDTKDLVCLGLSSIVGV
[•	ĺ	{	ĺ	i :	WYLLRKHWIANNLFGLAFSLNGVELLHLNN
'	1		1			VSTGCILLGGLFIYDVFWVFGTNVMVTVAKS
			l			FEAPIKLVFPQDLLEKGLEANNFAMLGLGDV
						VIPGIFIALLLRFDISLKKNTHTYFYTSFAAYIF
1	1	[GLGLTIFIMHIFKHAQPALLYLVPACIGFPVLV
Į.		1				ALAKGEVTEMFSYEESNPKDPAAVTESKEGT
						EASASKGLEKKEK
773	2123	A	6161	3	1088	CQPMLVTRKNHPKLLLRRTESVAEKMLTNW
			ŀ			FTFLLYKFLKESAGEPLFMLYCAIKHQMEKG
			[PIDAITGEARYSLSEDKLIRHLIDYKTLTLNCV
						NPENENAPEVPVKGLDCDTGTQAKEKLLDA
						AYKGVPYSQRPKAADMDLEWRQGRMARIIL
Ļ						QDEDVTTKIDNDWKRLNTLAHYQVTDGSSV
	ļ					ALVPKQTSAYNISNSSTFTKSLSRYESMLRTA
1						SSPDSLRSRTPMITPDLESGTKLWHLVKNHDH
			J .			LDQREGDRGSKMVSEIYLTRLLATKGTLQKF
				•		VDDLFETIFSTAHRGSALPLAIKYMFDFLDEO
]						ADKHQIHDADVRHTWKSNCLPLRFWVNVIK
						NPQFVFDIHKNSITDACLSVV
774	2124	A	6163	860	125	KTAVKKRNLNPVFNETLRYSVPQAELQGRVL
'''			****			SLSVWHRESLGRNIFLGEVEVPLDTWDWGSE
					1	PTWLPLQPRVPPSPDDLPSRGLLALSLKYVPA
						GSEGAGLPPSGELHFWVKEARDLLPLRAGSL
						DTYVQCFVLPDDSRASRQRTRVVRRSLSPVF
				,	l	NHTMVYDGFGPADLRQACAELSLWDHGALA
•				· .		NRQLGGTRLSLGTGSSYGLQVPWMDSTPEEK
775	2125				300	QLWQALLEQPCEWVDGLLPLRTNLAPRT
775	2125	Α	6191	2	392	ARGIGSLGRDHSGSGGGTGMAGAWVRKAAD
					ļ	YVRSKDFRDYLMSTHFWGPVANWGLPIAAIT
				l]	DMK\KSPEIISRRMTFAL*CYSLTFVRFAHYVQ
		'		ł	ľ	\PWNWLMLGCHTAVDFDQLISSMPCISHGMT
						ASASAL
776	2126	A	6217	1	827	FRGYWGVREAFTDASWSGGLGPGKPGMKIY
l				1		RQKHAKKHLGFFRNNFGVREPYQILLDGTFC
			İ		ľ	QAALRGRIQLREQLPRYLMGETQLCTTRCVL
						KELETLGKDLYGAKLIAQKCQVRNCPHFKNA
						VSGSECLLSMVEEGNPHHYFVATQDQNLSVK
						VKKKPGVPLMFIIQNTMVLDKPSPKTIAFVKA
		ĺ	ĺ	į		VESG\RLSQCMRKKVSNISKRNRV**KTLNRG
						RRKKRKKISGPNPLSCLKKKKKAPDTQSSASE
						KKRKRKRIRNRSNPKVLSEKQNAEGE
	ŀ					· · · · · · · · · · · · · · · · · · ·

..... erer samer takter samli Marija

SEQ ID SEQ ID Not of bod bod	GEO	TD CEOTID	Mot	LCEO	Predicted	Deadless and	I A mine solid seguence (A – Alemine C – Contains
				SEQ		Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
			1100				
Sequence			1	1			
Penice	l l	1 .	1				
minn aeid residue of peptide residue of peptide residue of peptide sequence peptide sequence peptide sequence per solo codon, pepsible nucleotide deletion, pepsible nucleotide deletion, pepsible nucleotide deletion, pepsible nucleotide deletion, pepsible nucleotide deletion, pepsible nucleotide deletion, pepsible nucleotide insertion nucleotide deletion, pepsible nucleotide deletion, pepsibl			1	1			
Peptide Sequence Peptide Sequence Peptide Sequence Peptide Sequence Peptide Sequence Peptide Sequence Peptide Sequence Peptide Sequence Peptid	uenc	٠		914			
		İ			1		
		1				sequence	
1777	1						
FOPFGEPRGGGHLRSGVILGGPGGKEPFFF YNSKISPALUGPPHYSLAGEAGGKEPFRC \text{RFQRGGAPLFSRVRGRAKIFLKKK \text{ASPFHHPIRGAFILLIARGS**GODGSLIHWSN \text{ASPSHHPIRGAFILLIARGS**GODGSLIHWSN \text{ASPADULDIKIN**I.DPLLEEKMPILEVKVVP \text{POVLSEPN**RSGGCFSAPSFEPPPHYTEGVEPKP \text{POVLSEPN**RSGGCFSAPSFEPPPHYTEGVEPKP \text{POVLSEPN**RSGGCFSAPSFEPPPHYTEGVEPKP \text{POVLSEPN**RSGGCFSAPSFEPPPHYTEGVEPKP \text{POVLSEPN**RSGGCFSAPSFEPPPHYTEGVEPKP \text{POVLSEPSN**RSGCFGCAPSGQDAGQBLKPRSGLGC \text{NSRYKRS \text{RAMBYSSSCRGGDAGGRELMGGIGKTM \text{MOSGGTGTMAIGMGRC**PWLPTTSVSPH \text{OSGMMY \text{RAMBYSSSCRGMRGRELMGGIGKTM \text{MOSGGTGTTMAIGMGRC**PWLPTTSVSPH \text{OSGMMY \text{RIMMCDRGIGMLITTVGAFAAFSLMTIAVG \text{TDYWLYSRGVCRTKSTSIDMETSRNEEWMT \text{MOSGGTTGTTMAIGMGRC**PWLPTTSVSPH \text{OSGMMY \text{CSGMMY \text{RIMMCDRGIGMLITTVGAFAAFSLMTIAVG \text{TDYWLYSRGVCRTKSTSIDMETSRRNEEWMT \text{HSGWCRTCLEGAFGVCKKIDHFEDADYE \text{QDTAEYLLRAVRASSYPPLSVTLLFFGGLCV \text{AASEFFRSHNVLSAGFFVSAGLSINGIGIVYI \text{SANAGRTPGGROSSKKSYSYGWSFYTSGAFS \text{PRIMMCTCLEGAFGVCKKIDHFEDADYE \text{QDTAEYLLRAVRASSYPPLSVTLLFFGGLCV \text{AASEFFRSHNVLSAGFFVSAGLSHIGITYSICH \text{STFARLPYPXPRRRSSSSSSTEPRSRDLS \text{PRIMMCTCLEGAFGVCKKIDHFEDADYE \text{QDTAEYLLRAVRASSYPPLSVTLLFFGGLCV \text{AASEFFRSHNVLSAGFFVSAGLSHIGITYSICH \text{STFARLPYPXPRRRSSSSSSTEPRSRDLS \text{PRIMMCTCLEGAFGVMSFYTLSFGAFCVCKIDHFEDADYE \text{QDTAEYLLRAVRASSYPPLSVTLLFFGGLCV \text{AASEFFRSHNVLSAGFFVSAGLSHIGHTSTICLSWGGOL \text{LEGAGGAFGVAGAFSNSAGFFSTSCAGLSHIGHTSTICLSWGGOL \text{LEGAGGAFGVAGAFSNSAGFFSTSCAGLSHIGHTSTICLSWGGOL \text{LEGAGGAFGVAGAFSNSAGFFSTSPCAGLS \text{MOSGAFGLAGGAFGAFGAGAFAGAFSNGAAFSLASSNKKNSKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK							
	1777	2127	A	6236	1038	1402	YYQISSLPSIVGNGIFLWLLICIFLAKQGGSRL*
RPQRGGIAPLPSRYRGRAKIFLKKK		[1		f	[
778		ļ					
AVSNADILIDIKIN*IDIRILERMPILERKWPILERW SPORDOGAL GOOPPLOIPSDSILALLKKOT*RA LINWPLOSIARRSSCFGOOPODILKPRSOLOC SPORDOGAL GOOPPLOIPSDSILALLKKOT*RA LINWPLOSIARRSSCFGOOPODILKPRSOLOC SPORDOGAL GOOPPLOIPSDSILALLKKOT*RA LINWPLOSIARRSSCFGOOPODILKPRSOLOC SPORT VR	L						
PQVI.SEPN'RSGCCSAPSEVPPWTGEVKP SPQRDGGALGGQCG JGIPSDISLALKKQI'RA	778	2128	A	6237	422	913	ASFFHHHRGAFLLLLAIPGS*GQDQSLIHWSN
SPÖRDGALGQGPLGIPSDSILALLKKQT*RA	1	ì	i		ĺ	1	AVSNAD\LLDLK\N*LDH\LEEKMPL\EVKVVP
LLNWPLGSLRRSSCFGGQDGQDLKPRSGLGC NSFRYRR							PQVL\SEPN*RSGGCFSAPSFEVPPWTGEVKP/
NSFRYRR		i					SPQRDGGALG\QGPLGIPSDSILALLKKQT*RA
779							LLNWPLGSLRRSSCFGGQDGQDLKPRSGLGC
YGQSQPSCTDRVKMGFVMGCAVGMAAGAL FGTFSCLSSILVSSSG/SGMRGELMGGIGKTM MQSGGTTGTTFMAIGMGIRC*PWLPTTSVPSH QSQPMY							NSFRYRR
YGQSQPSCEDRVKMGFVMGCAVGMAGAL FGTFSCLSSLI VSSSGYGKRELMGGIGKTM FGTFSCLSSLI VSSSGYGKRELMGGIGKTM MQSGTFGTFMAIGMGIRC*PWLPTTSVPSH QSQPMY SRMCRITGAFAFSLMTIAVG TDYWLYSRGVCRTKSTSDNETSRRNEEVMT HSGLWRTCCLEGAFRGVKIBHPPEDADYE QDTAEYLLRAVRASSVPILSVTILJFGGLCV AASSFRRSRHNVILSAGTFVSAGLSNIIGITVJ SAMAGRTPQRIDSKKSYSYGWSF/YFSGAFS FIIGRIIC*GVGLPWHIYIEKHQLEAKSHSEF LKKSTFARLPPYRYFRFRSSSTEPRSRDLS PISKGFHTIPSTDISMFTLSRDPSKITTMGTLLNS DRDHAFLQFHNSTPKEFKESLHNPANRRIT PV PROMOGOMVALFQDSIAWSNKSNPSSTSPRSRDLS PISKGFHTIPSTDISMFTLSRDPSKITTMGTLLNS QSPCQVQAPEGPSSHLPTLSFTTCLSWQGGD LEFLGDLKGCSELKNFQELTCSALVHPKADV WYCGRPLGFISSTHCHFTSTLSTTCLSWQGGD LEFLGDLKGCSELKNFQELTTQSALVHPKADV WYCGRPLGTLSSN WISLPSSLLCRKNGSSAEDDRRIGEPSALEAG EREDWGIGSA*SVGAVSKYPSARP*RTYPSE DEEEVTHQKSSSSDSNSEHRKKKTSKKNKKTKK KKRKNKSSKRKHRKKYSDSDSNSESDTNSDSD DDKKRYKKAKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK	779	2129	A	6249	420	36	ARAPSPSFSVRDVELSDPARERGEMPVAVGP
FGTTSCLSSILVSSSG/SGMRGRELMGGIGKTM MQSGGTFGTMAIGMGIRC*PWLPTTSVPSH QSQPMY			1				
MQSGGTFGTFMAIGMGIRC*PWLPTTSVPSH QSQPMY							
		1					
780	•						
TDYWLYSROVCRTK STSUNETSRENEEVMT HSGLWRTCCLEGAFRGVCKKIDHFPEDADYE QDTAEYLLRAVRASSYPHLSVTLLFFGGLCV AASEFHRSRINVILSAGIFFVSAGLSNIIGIVYTI SANAGRFROGROSKSYSV GWSPYNFSGAFS FIIGRIIC*GVGLPWHIYIEKHQQLRAKSHSEF LKKSTFARLPPYRYRRRRSSSRTEPRSRDLS PISKGHTHTSTDISMFTISRDPSKITMGTILINS DRDHAFLQFHNSTPKEFKESLHNNPANRRTT PV 781 2131 A 6274 832 318 RIIKVKDLKQTLAIKTAYPRCKCLVEMDQIFH LQVKQKQLACLCTWQARDPDCPSTKVVLL VGPGMGCMVALFQDSIAWSNKSMPSSLSAIS QSPCQVQAPEGPSSFHLPTLSFTTCLSWQGGD LEFLGDLKGCSELKNFQELITQSALVHPKADV WYVCGRPLLGTLPSN WWYCGRPLLGTLPSN REPUBLICATION WYVCGRPLLGTLPSN REPUBLICATION TO BE TO	780	2130	Δ	6263	415	1380	
HSGLWRTCCLEGAFRGVCKKIDHFPEDADYE QDTAEYLLRAVRASSVFPILSVTLLFFGGLCV AASEFHRSRINVILSAGIFFVSAGLSNIIGIIVYI SIANAGRIFQGRIDSKKSYSYGWSFYFSGAFS FIIGRIIC-GVGLPWHYIEKHQQLRAKSHSEF LKKSTFARLPYNRYRFRRSSSRSTEPRSRDLS PISKGHTIPSTDISMFTLSRDPSKITMGTILNS DRDHAFLQFINSTPIKEFKESLINNNPANRRITT PV 781 2131 A 6274 832 318 RIIKVKDLKQTLAIKTAYPRCKCLVEMDQIFH LQVKQKQLACLCTWQARDPDCPPSTKVVL/L VGPOMGCMYALFQDSIAWSNKSMPSSLSAIS QSPCQVQAPEGPSSFHLPTLSFTTCLSWQGDD LEFLGDLKGCSELKNRQELITQSALVHFKADV WYVCGRPLGTLPSN WYVCGRPLGTLPSN WYVCGRPLGTLPSN WISLPSSLLCRKNGSSAEDDRR\GEPSAEEAEG EREDWGIGSA*SVGAVSKYPSAFF*RTYPSE DEEVTHOKSSSSDSNSSEHRKKKTSRSNK KKRNKSSKRKHRYYSDSDONSSEDHRKKKTKK ESSDSSCKDSEEDLSEATWMEQPNVADTMDL IGFEAPIIHTSQDEKPLKYGHALLPGEGAAMA EYYKAGKHPRGEGILSEGISFECSGYVM SGSRHRMEAVRLRKENQIYSADEKRALASF NQEERRRESKLLASFREMYHKKTKKKCDNK WDDYPQGALRREAAEGLHFLGPFGRVRQQ LRGITGPAWYCHSPSHSLLSAFCHLPTPSRCP AMARPPYGSVVVPNWESRRRQGQVSPGLBS AQEPPAGVWAA*AASAAAALSIDTASYKIFV SGKSGVGKTALVAKLAGLEVPVVHHETTIGIQ TTVVTWPAKLQASSRVVMFRIEFWDCGESA LKKPDHMLLACMENTDAFLFLFSFDRASFE DLFGGLARIAGEAPGVVRMVIGSKFDGYMHT DVPERDLTAFRQAWELPLLRVKSVPGRRLG SSPPASLITMKNQDKKNGAKQSNPKSP GQFAAGFEGAQARTAQSGALRDVSSEP GGSPPASLITMKNQDKKNGAKGSNPKSP GQPEAGFEGAQARTAQSGAARAVEAEGFGSSQA PRKPEGAQARTAQSGALRDVSSEPSKQDPNITEIR QSDEVGDRDHRRPQEKKKAGLIGKEITLLM QTLNILSTFEEKLAALCKKYABLLEEHENSQ	1,00	2130	1 ^	0203	413	1360	
ODTAEVILRAVRASSVPPILSVTLLFFGGI_CV AASEFHRSRHNVILSAGIFVSAGLSNIIGIIVYI SANAAGRIPQQRIDSKKSYSYGWSPYFTSGAFS FIGR/IIC*GVGLPWHIYIEKHQQLRAKSHSEF LKKSTFARLPPYRYRPRRSSSRSTEPSRDLS PISKGFHTIPSTDISMFTLSRDPSKITMGTILNS DRDHAFLQFHNSTPKEFKESLHNNPANRRIT PV 781 2131 A 6274 832 318 RIIKVKDLKQTLAIKTAYPRCKCLVEMDQIFH LQVKQKQLACLCTWQARDPDCPPSTKVVL/L VGPGMGCMVALFQDSIAWSNKSMPSSLSAIS QSPCQVQAPEGPSSFHLPTLSFTTCLSWQGGD LEFLGDLKGCSELKHNPGLITQSALVHPKADV WWYCGRPLLGTLPSN 782 2132 A 6281 1324 393 WISLPSSLLCKKNGSSAEDDRRGEPSAEEAE EREDWGIGSA*SVGAVSKVPSAR*PRTYPSE DEEEVTHQKSSSSDSNSEEFIRKKKTSRSRNK KKRKNKSSKRKHRKYSDSDSNSESDTNSDSD DDKKRVKAKKKKKKHTKKKKNKKTKK ESSDSSCKDSEEDLSEATWMEOPNVADTMDL IGFEAPIHTSQDEFLKYGHALJEGGAAMA EYVKAGKRIPRGEIGLTSEEIGSFECSGYVM SGSRHRRMEAVRLRKENQIYSADEKRALASF NQEERKRESKILASFREMVHKKTKGKDDK AQEPPAGVWAA*AASAAAALSIDTASYKIFV SGKSGVGKTALVAKLAGLEVPVVHHETTGIQ TTVVFWPAKLQASSRVVMFREFWDCGESA LKKPDHMLACMENTDAFLFFFTDRASFE DLFQQLARIAGEAPGVVRMVIGSKFDQYMHT DVPERDLTAFRQAWELPLRVKSVPGRRLG 784 2134 A 6308 86 96 GSSPDPASLTIMKNOKKNSKRGKNSFSSQA PRKPEGAQARTAQGALRDVSELSRQLEDIL STYCVDNNQGGGEGROAGEPASPEDAEKSR TYVARNGEPEPTPVVNGEKEPSKGDPNTEEIR QSDEVGDRDHRRPQEKKKAKGLGKEITLLM QTINTLSTTPEEKLAALCKKYAELLEEHRNSQ	1			1			
ASSEHRSRHNVILSAGIFFVSAGLSNIIGIIVYI SANAGRTPGQNDSKKSYSYGWSF/YFSGAFS FIGR/IIC*GVGLPWHIVIEKHQQLRAKSHSEF LKKSTFARLPYRYRFRRKSSRSTEPRSRDLS PISKGFHTIPSTDIBMT1SRDPSKITMGTLLNS DRDHAFLQFHNSTPKEFKESLHNNPANRRTT PV 781 2131 A 6274 832 318 RIIKVKDLKQTLAIKTAYPRCKCLVEMDQIFH LQVKQKQLACLCTWQARDPDCPPSTKVVL/L VGPGMGCMVALFQDSIAWSNKSMPSSLSAIS QSPCQVQAPEGPSSLPHTISFTTCLSWQGDD LEFLGDLKGCSELKNFQELITQSALVHPKADV WWYCGRPLLGTLPSN 782 2132 A 6281 1324 393 WISLPSSLLCRKNGSSAEDDRRIGEPSAEEAEG EREDWGIGSA*SVGAVSKVYSARF*RTYPSE DEEVTHQKSSSSDSNSEEHRKKKTSRSRNK KKRKNKSSKRKHRKYSDSDNSSESDTNSDSD DDKKRVKAKKKKKKKKHKTKKKKSNKKTKK ESSDSSCKDSEEDLSATMBEQPNVADTMDL IGPEAPIHTSQDEKPLKYGHALLPGEGAAMA EYVKAGKRIPRRGEIGLTSEEIGSFECSGYVM SGSRHRRMEAVRLRKENQIYSADEKRALASF NQEERRKRESKILASREMVHKKTKGKDDK 783 2133 A 6305 201 1032 WDDYPQGALRRREAAEGHFLGPFGRVRGQ LRGITGPAWYCHSPSHSLLSAFCHLPTPSRCP AMARPPYPGSVVVPNWHESRRGQGVPGLHS AQEPPAGWWAA*ASAAAAUSIDTASYKIFV SGKSGVGKTALVAKLAGLEVPVVHHETTGIQ TTVVFWPAKLQASSRVVMRFEFWDCCESA LKKPDHMLLACMENTDAFLFLSFPTDRASFE DLPGQLARIAGEAPQVVRMVIGSKFDQYMHT DVPERDLTAFRQAWELPLLRVKSVPGRRLG TVVFWPAKLQASSRVVMRFEFWDCCESA LKKPDHMLLACMENTDAFLFLSFTPTRASFE DLPGQLARIAGEAPGVVRMVIGSKFDQYMHT DVPERDLTAFRQAWELPLLRVKSVPGRRLG GSSPDPASLITMKNOKNINGAGARGVPGSSQA PRKPEGAQAGRTAQSALARDVSELSRQLEDIL STYCVDNNQGGFGEDGAQGEPAPEDAEKSR TYVARNGEPEPTPVVNGEKEPSKGDPNTEEIR QSDEVGDRDHRRPQEKKKAKGLGKEITLLM QTINTILSTTPEEKLAALCKKYAELLEEHRNSQ							I
SANAGRITGORIDSKKSYSYGWSFY/TSGAFS FIIGR/IIC+GVGLPWHIYIEKHQQLRAKSHSEF LKKSTFARLPPYRYRFRRRSSSRSTEPRSDLS PISKGFHITIPSTDISMFTLSRDPSKITMGTLLNS DRDHAFLQFHNSTPKEFKESLHNNPANRATT PV 781 2131 A 6274 832 318 RIIKVKDLKQTILAKTAYPRCKCLVEMDQIFH LQVKQKQLACLCTWQARDPDCPPSTKVVL/L VGPGMGGMVALFQDSIAWSNKSMPSSLSAIS QSPCQVQAPEGPSSFHLPTLSFTTCLSWQGGD LEFLGDLKGCSELKNFQELITQSALVHPKADV WWYCGRPLLGTLPSN 782 2132 A 6281 1324 393 WISLPSSLLCRKNGSAEDDRRIGEPSAEEAEG EREDWGIGSA*SVGAVSKVPSARF*RTYPS\(\text{E}\) DEEEVTHQKSSSSDSNSEEHRKKTTSRSTNK KRKNKSSKRKHRKYSDSDSNSESDTNSDSD DDKKRVKAKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK							1 •
FIGGRIIC*GVČLPWHYJEKHQQLRAKSHSEF LKKSTPARLPYYRYRPRRSSSRSTEPRSRDLS PISKGFHTIPSTDISMFTLSRDPSKITMGTLLNS DRDHAFLQFHNSTPKEFKESLHNNPANRRTT PV 781 2131 A 6274 832 318 RIIKVKDLKQTLAIKTAYPRCKCLVEMDQIFH LQVKQKQLACL.CTWQAADPDCPPSTKVVL/L VQPGMGCMVALFQDSIAWSNKSMPSSLSAIS QSPCQVQAPEGPSSFHLPTLSFTTCLSWQGGD LEFI.GDLKGCSELKNFQELITQSALVHPKADV WWYCGRPLLGTLPSN 782 2132 A 6281 1324 393 WISLPSSLLCRKNGSSAEDDRRIGEPSAEEAEG EREDWGIGSA*SVGAVSKVPSARF*RTYPSSE DEEEVTHQKSSSDSNSEBERKKKTSRSRNK KRKNKSSSRSMSNSEHRKKKTSKSRNK KRKNKSSKRKHRKYSDDSNSESDTNSDSD DDKKRVKAKKKKKKKKKKKKKKKKKKKK ESSDSSCXDSEEDLSEATWMEQPPVADTMDL IGPEAPHHTSQDEKPLKYGHALLPGEGAAMA EYVKAGKRIPRRGEIGLTSEEIGSFECSGYVM SGSRHRMEAVRLRKENQIVSADEKRALASF NQEERKKRESKILASFREWYHKKTKGKDDK 783 2133 A 6305 201 1032 WDDYPQGALRREAAEGLHFLGPPGRVRQQ LRGTIGPAWYCHSPSHSLLSAFCHLPTTSRCP AMARPPVPGSVVVPNWHES/RRQQGVPGLHS AQEPPAGVWAA*AASAAALSIDTASYKIFV SGKSGVGKTALVAKLAGLEVPVVHHETTGIQ TTVVFWPAKLQASSRVVWNFREFEWDCGESA LKKFDHMLLACMENTDAFLFIFSPTDRASFE DLPGQLARIAGEAPGVVRWVIGSKFDQYMHT DVPERALLARGEAPGVVRWVIGSKFDQYMHT DVPERALLARGEAPGVVRWVIGSKFDQYMHT DVPERALLARGEAPGVVRWVIGSKFDQYMHT DVPERALLARGEAPGSQA PREPEGAQARTAQSGALROVSEELSRQLEDIL STYCVDNNQGGPGEDGAQGEPAEPEDAEKSR TYVARNGEPEPTPVVNGEKEPSKGDPNTEEIR QSDEVODRDHRRPQEKKKAKGLIKETILLM QTLNTLSTFEEKLAALCKKYABELLEEHRNSQ		1	1	[
LKKSTFARLPPYRYRPRRSSSSTEPRSRDLS PISKGFHTIPSTDISMFTLSRDPSKITMGTLLNS DRDHAFLQFHINSTYKEFKESLHNNPANRRTT PV		ĺ					
PISKGFHTPSTDISMFTLSRDPSKITMGTLLNS			-	1			
781 2131 A 6274 832 318 RIIKVKDLKQTLAIKTAYPRCKCLVEMDQIFH LQVKQKQLACLCTWQARDPDCPPSTKYVLLL VGPGMGCMVALFQDSIAWSNKSMPSSLSAIS QSPCQVQAPEGPSSFHLPTLSFTTCLSWQGGD LEFLGDLKGCSELKNFQELITQSALVHPKADV WWYCGRPLLGTLPSN 782 2132 A 6281 1324 393 WISLPSSLLCRKNGSSAEDDRR\GEPSAEEAEG EREDWGIGSA*SVGAVSKVPSARF*RTYPS\(\text{E}\) DEEVTHQKSSSSDSNSEEHRKKKTSRSRK KKKRKNKSSKKHFRKYSDSDSNSESDTNSDSD DDKKRVKAKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK			1	i			
781 2131 A 6274 832 318 RIIKVKDLKQTLAIKTAYPRCKCLVEMDQIFH LQVKQKQLACLCTWQARDPDCPPSTK.VVL/L VGPGMGCMV.ALFQDSIAWSNKSMPSSLSAIS QSPCQVQAPEGPSSFH.PILSFITTCLSWQGGD LEFLGDLKGCSELKNFQELITQSALVHPKADV WWYCGRPLLGTLPSALVHPKADV WWYCGRPLLGTLPSN WWYCGRPLLGTLPSN WISLPSSLLCRKNGSSAGDDRRKGEPSAEEAEG EREDWGIGSA*SVQAVSKVPSARF*RTYPSIE DEEVTHQKSSSSDSSSEDSEBEHRKKKTSRSRNK KKRKNKSSKRKHRKYSDSDSNSESDTNSDSD DDKKKVKAKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK				1			PISKGFHTIPSTDISMFTLSRDPSKITMGTLLNS
781]			j			DRDHAFLQFHNSTPKEFKESLHNNPANRRTT
LQVKQKQLACLCTWQARDPDCPPSTKVUL/L VGPGMGCMVALFQDSIAWSNKSMPSSLSAIS QSPCQVQAPEGPSSFHI.PTLSFTTCLSWQGGD LEFLGDLKGCSELKNFQELITQSALVHPKADV WWYCGRPLLGTLPSN 782 2132 A 6281 1324 393 WISLPSSLLCKKNGSSAEDDRR\GEPSAEEAEG EREDWGIGSA*SVGAVSKVPSARF*RTYPS\(\text{E}\) DEEEVTHQKSSSDSNSEEHRKKKTSRSRNK KKRKNKSSKRKHRKYSDSDSNSEDTNSDSD DDKRRVKAKKKKKKKKKKKKKKTKKKNKTKK ESSDSSCKDSEEDLSEATWMEQPNVADTMDL IGPEAPIHTSQDEKPLKYGHALLPGEGAAMA EYVKAGKRIPRRGEIGLTSEEIGSPECSGYVM SGSRHRRMEAVRLRKENQIYSADEKRALASF NQEERRKRESKILASFREMVHKKTKGKDDK 783 2133 A 6305 201 1032 WDDYPQGALRREAAEGLHFLGPPGRVRGQ LRGTTGPAWYCHSPSHSLLSAFCHLPTPSRCP AMARPVOGSVVYPNWHES/RRGQGVPGLHS AQEPPAGVWAA*AASAAAALISIDTASYKIFV SGKSGVGKTALVAKLAGLEVPVVHHETTGIQ TTVVFWPAKLQASSRVVMFRFEFWDCGESA LKKFPHMLLACMENTDAFLFLFSFTDRASFE DLPGQLARIAGEAPGVVRMVIGSKFDQYMHT DVPERDLTAFRQAWELPLRVKSVPGRRLG 784 2134 A 6308 86 96 GSSPDPASLITIMKNQDKKNGAAKQSNPKSSP GQPEAGPEGAQERPSQAAPAVEAEGPGSSQA PRKPEGAQARTAQSGALRDVSEELSRQLEDLL STYCVDNNQGGPGEDGAQGEPAEPEDAEKSR TYYARNGEPEPTPVVNGEKEPSKGDPNTEEIR QSDEVGROPHTPOVKKKKAKGLGLKEGTLLM QTLNTLSTPEEKLAALCKKYAELLEEHRNSQ							PV
LQVKQKQLACLCTWQARDPDCPPSTKVUL/L VGPGMGCMVALFQDSIAWSNKSMPSSLSAIS QSPCQVQAPEGPSSFHI.PTLSFTTCLSWQGGD LEFLGDLKGCSELKNFQELITQSALVHPKADV WWYCGRPLLGTLPSN 782 2132 A 6281 1324 393 WISLPSSLLCKKNGSSAEDDRR\GEPSAEEAEG EREDWGIGSA*SVGAVSKVPSARF*RTYPS\(\text{E}\) DEEEVTHQKSSSDSNSEEHRKKKTSRSRNK KKRKNKSSKRKHRKYSDSDSNSEDTNSDSD DDKRRVKAKKKKKKKKKKKKKKTKKKNKTKK ESSDSSCKDSEEDLSEATWMEQPNVADTMDL IGPEAPIHTSQDEKPLKYGHALLPGEGAAMA EYVKAGKRIPRRGEIGLTSEEIGSPECSGYVM SGSRHRRMEAVRLRKENQIYSADEKRALASF NQEERRKRESKILASFREMVHKKTKGKDDK 783 2133 A 6305 201 1032 WDDYPQGALRREAAEGLHFLGPPGRVRGQ LRGTTGPAWYCHSPSHSLLSAFCHLPTPSRCP AMARPVOGSVVYPNWHES/RRGQGVPGLHS AQEPPAGVWAA*AASAAAALISIDTASYKIFV SGKSGVGKTALVAKLAGLEVPVVHHETTGIQ TTVVFWPAKLQASSRVVMFRFEFWDCGESA LKKFPHMLLACMENTDAFLFLFSFTDRASFE DLPGQLARIAGEAPGVVRMVIGSKFDQYMHT DVPERDLTAFRQAWELPLRVKSVPGRRLG 784 2134 A 6308 86 96 GSSPDPASLITIMKNQDKKNGAAKQSNPKSSP GQPEAGPEGAQERPSQAAPAVEAEGPGSSQA PRKPEGAQARTAQSGALRDVSEELSRQLEDLL STYCVDNNQGGPGEDGAQGEPAEPEDAEKSR TYYARNGEPEPTPVVNGEKEPSKGDPNTEEIR QSDEVGROPHTPOVKKKKAKGLGLKEGTLLM QTLNTLSTPEEKLAALCKKYAELLEEHRNSQ	781	2131	A	6274	832	318	RIIKVKDLKQTLAIKTAYPRCKCLVEMDQIFH
OSPCQVQAPEGPSFHILPTLSFTTCLSWQGGD LEFLGDLKKGCSELKNFQELITQSALVHPKADV WYYCGRPLLGTLPSN 782 2132 A 6281 1324 393 WISLPSSLLCRKNGSSAEDDRR\GEPSAEEAEG EREDWGIGSA*SVGAVSKVPSARF*RTYPS\(\text{E}\) DEEVTHQKSSSSDSNSEEHRKKTSRSRNK KRKNKSSKRKHRKYSDSDSNSESDTNSDSD DDKKRVKAKKKKKKKKKKKKKKKKKKKKKKK ESSDSSCKDSEEDLSEATWMEQPNVADTMDL IGPEAPIHTSQDEKPLKYGHALLPGEGAAMA EYVKAGKRIPRRGEIGLTSEEIGSFECSGYVM SGSRIRRMEAVRLRKENQIYSADEKRALASF NQEERRKRESKILASFREMVHKKTKGKDDK 783 2133 A 6305 201 1032 WDDYPQGALRREAAEGLHFLGPPGRVRGQ LRGITGPAWYCHSPSHSLLSAFCHLPTPSRCP AMARPPYPGSVVVPNWHES/RRGQGYPGLHS AQEPPAGVWAA*AASAAAA\SIDTASYKIFV SGKSGVGKTALVAKLAGLEYPVVHHETTGIQ TTVVFWPAKLQASSRVVMFRFEFWDCGESA LKKFDHMLLACMENTDAFLFLFSFTDRASFE DLPGQLARIAGEAPGVVRMVIGSKFDQYMHT DVPERDLTAFRQAWELPLLRVKSVPGRRLG 784 2134 A 6308 86 96 GSSPDPASLITMKNQDKKNGAAKQSNPKSSP GQPEAGPEGAQERPSQAAPAVEAEGPGSSQA PRKPEGAQARTAQSGALRDVSEELSRQLEDIL STYCVDNNQGGPGEDGAQGEPAEPEDAEKSR TYVARNGEPEPTIPVVNGEKEPSKGDPNTEEIR QSDEVGDRDHRRPQEKKKAKGLGKEITLLM QTLNTLSTPEEKLAALCKKYAELLEEHRNSQ		1					
OSPCQVQAPEGPSFHILPTLSFTTCLSWQGGD LEFLGDLKKGCSELKNFQELITQSALVHPKADV WYYCGRPLLGTLPSN 782 2132 A 6281 1324 393 WISLPSSLLCRKNGSSAEDDRR\GEPSAEEAEG EREDWGIGSA*SVGAVSKVPSARF*RTYPS\(\text{E}\) DEEVTHQKSSSSDSNSEEHRKKTSRSRNK KRKNKSSKRKHRKYSDSDSNSESDTNSDSD DDKKRVKAKKKKKKKKKKKKKKKKKKKKKKK ESSDSSCKDSEEDLSEATWMEQPNVADTMDL IGPEAPIHTSQDEKPLKYGHALLPGEGAAMA EYVKAGKRIPRRGEIGLTSEEIGSFECSGYVM SGSRIRRMEAVRLRKENQIYSADEKRALASF NQEERRKRESKILASFREMVHKKTKGKDDK 783 2133 A 6305 201 1032 WDDYPQGALRREAAEGLHFLGPPGRVRGQ LRGITGPAWYCHSPSHSLLSAFCHLPTPSRCP AMARPPYPGSVVVPNWHES/RRGQGYPGLHS AQEPPAGVWAA*AASAAAA\SIDTASYKIFV SGKSGVGKTALVAKLAGLEYPVVHHETTGIQ TTVVFWPAKLQASSRVVMFRFEFWDCGESA LKKFDHMLLACMENTDAFLFLFSFTDRASFE DLPGQLARIAGEAPGVVRMVIGSKFDQYMHT DVPERDLTAFRQAWELPLLRVKSVPGRRLG 784 2134 A 6308 86 96 GSSPDPASLITMKNQDKKNGAAKQSNPKSSP GQPEAGPEGAQERPSQAAPAVEAEGPGSSQA PRKPEGAQARTAQSGALRDVSEELSRQLEDIL STYCVDNNQGGPGEDGAQGEPAEPEDAEKSR TYVARNGEPEPTIPVVNGEKEPSKGDPNTEEIR QSDEVGDRDHRRPQEKKKAKGLGKEITLLM QTLNTLSTPEEKLAALCKKYAELLEEHRNSQ			1'	1			
TEFLGDLKGCSELKNFQELITQSALVHPKADV WWYCGRPLLGTLPSN 782 2132 A 6281 1324 393 WISLPSSLCRKNGSSAEDDRRIGEPSAEEAEG EREDWGIGSA*SVGAVSKVPSARF*RTYPSIE DEEEVTHQKSSSSDSNSEEHRKKKTSRSRNK KKRKNKSSSKRKHRK YSDSDSNSESDTNSDSD DDKKRVKAKKKKKKKKKKKKKKKKK ESSDSSCKDSEEDLSEATWMEQPNVADTMDL IGPEAPIIHTSQDEKPLKYGHALLPGEGAAMA EYVKAGKRIPRGEGIGLTSEEIGSFECSGYVM SGSRHRMEAVPLREENQIVSADEKRALASF NQEERKKESKILASFREMVHKKTKGKDDK 783 2133 A 6305 201 1032 WDDYPQGALRREAAEGLHFLGPPGRVRGQ LRGITGPAWYCHSPSHSLLSAFCHLPTPSRCP AMAPPYPGSVVVPNWHES/RRGQGVPGLHS AQEPPAGVWAA*AASAAAALSIDTASYKIFV SGKSGVGKTALVAKLAGLEVPVVHHETTGIQ TTVVFWPAKLQASSRVVMFRFEFWDCGESA LKKFDHMLLACMENTDAFLFLFSPTDRASFE DLPGQLARIAGEAPGVVRMVIGSKFDQYMHT DVPERDLTAFRQA WELPLLRVKSVPGRRLG 784 2134 A 6308 86 96 GSSPDPASLITMKNQDKKNGAAKQSNPKSSP GOPEAGPEGAQERPSQAAPAVEAEGPGSSQA PRKPEGAQARTAQSGALRDVSEELSRQLEDIL STYCVDNNQGGPGEDGAQGEPAEPEDAEKSR TYVARNGEPEPIPVVNGEKEPSKGDPNTEEIR QSDEVOGRDHRRPQEKKKAKGLGKEITLLM QTLNTLSTPEEKLAALCKKYAELLEEHRNSQ	i	1	ľ	1			
782 2132 A 6281 1324 393 WISLPSSLLCRKNGSSAEDDRR\GEPSAEEAEG EREDWGIGSA*SVGAVSKVPSARF*RTYPS\E DEEEVTHQKSSSSDSNSEEHRKKKTSRSRNK KKRKNKSSKRKHRKYSDSDSNSESDTNSDSD DDKKRVKAKKKKKKKHKTKKKNKKTKK ESSDSSCKDSEDLSEATWMEQPNVADTMDL IGPEAPIHTSQDEKPLKYGHALLPGEGAAMA EYVKAGKRIPRRGEIGLTSEEIGSFECSGYVM SGSRHRRMEAVRLRKENQIYSADEKRALASF NQEERRKRESKILASFREMVHKKTKGKDDK 783 2133 A 6305 201 1032 WDDYPQGALRREAAEGLHFLGPPGRVRGQ LRGITGPAWYCHSPSHSLLSAFCHLPTPSRCP AMARPPVPGSVVVPNWHES/RRGQGVPGLHS AQEPPAGVWAA*AASAAAALSIDTASYKIFV SGKSGVGKTALVAKLAGLEVPVVHHETTGIQ TTVVFWPAKLQASSRVVMFRFEFWDCGESA LKKFDHMLLACMENTDAFLFLFSFTDRASFE DLPGQLARIAGEAPGVVRMVIGSKFDQYMHT DVPERDLTAFRQAWELPLLRVKSVPGRRLG 784 2134 A 6308 86 96 GSSPDPASLITMKNQDKKNGAAKQSNPKSSP GQPEAGPEGAQERPSQAAPAVEAEGPGSSQA PRKPEGAQARTAQSGALRDVSEELSRQLEDIL STYCVDNNQGGPGEDGAQGEPAEPEDAEKSR TYVARNGEPEPTPVVNGEKEPSKGDPNTEEIR QSDEVGDRDHRRPQEKKKAKGLGKETTLLM QTLNTLSTPEEKLAALCKKYAELLEEHRNSQ			į		•		
782 2132 A 6281 1324 393 WISLPSSLLCRKNGSSAEDDRR\GEPSAEEAEG EREDWGIGSA*SVGAVSKVPSARF*RTYPS\(\text{PSAEEAEG}\) EREDWGIGSA*SVGAVSKVPSARF*RTYPS\(\text{PSAEEAEG}\) EREDWGIGSA*SVGAVSKVPSARF*RTYPS\(\text{PSAEEAEG}\) DEEVTHQKSSSSDSNSESDTNSDSD DDKKRVKAKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK				1			
EREDWGIGSA*SVGAVSKVPSARF*RTYPSVE DEEVTHQKSSSSDSNSEEHRKKKTSRSNKK KKRKKSSKRKHRKYSDSDSNSESDTNSDSD DDKKRVKAKKKKKKKKKKKKKKKKKKKKKKKKK ESSDSSCKDSEEDLSEATWMEQPNVADTMDL IGPEAPIHTSQDEKPLKYGHALLPGEGAAMA EYVKAGKRIPRRGEIGLTSEEIGSFECSGYVM SGSRHRREAVRLRKENQIYSADEKRALASF NQEERRKRESKILASFREMVHKKTKGKDDK 783 2133 A 6305 201 1032 WDDYPQGALRREAAEGLHFLGPPGRVRGQ LRGITGPAWYCHSPSHSLLSAFCHLPTPSRCP AMARPPVPGSVVVPNWHES/RRGQGVPGLHS AQEPPAGVWAA*AASAAAALSIDTASYKIFV SGKSGVGKTALVAKLAGLEVPVVHHETTGIQ TTVVFWPAKLQASSRVVMFRFEFWDCGESA LKKFDHMLLACMENTDAFLFLFSFTDRASFE DLPGQLARIAGEAPGVVRMVIGSKFDQYMHT DVPERDLTAFRQAWELPLLRVKSVPGRRLG 784 2134 A 6308 86 96 GSSPDPASLITMKNQDKKNGAAKQSNPKSSP GQPEAGPEGAQARTAQSGALRDVSEELSRQLEDIL STYCVDNNQGGPGEDGAQGEPAEPEDAEKSR TYVARNGEPEPIPVVNGEKEPSKGDPNTEEIR QSDEVGDRDHRRPQEKKKAKGLGKEITLLM QTLNTLSTPEEKLAALCKKYAELLEEHRNSQ	782	2132	A -	6281	1324	303	
DEEEVTHQKSSSSDSNSEEHRKKKTSRSNK KKRKNKSSKRKHRKYSDSDSNSESDTNSDSD DDKKRVKAKKKKKKKKKKKKKKKKKK ESSDSSCKDSEEDLSEATWMEQPNVADTMDL IGPEAPIIHTSQDEKPLKYGHALLPGEGAAMA EYVKAGKRIPRRGEIGLTSEEIGSFECSGYVM SGSRHRMEAVRLRKENQIYSADEKRALASF NQEERKKESKILASFREMVHKKTKGKDDK 783 2133 A 6305 201 1032 WDDYPQGALRREAAEGLHFLGPPGRVRGQ LRGITGPAWYCHSPSHSLLSAFCHLPTPSRCP AMARPPVPGSVVVPNWHES/RRGQGVPGLHS AQEPPAGVWAA*AASAAAA\LSIDTASYKIFV SGKSGVGKTALVAKLAGLEVPVVHHETTGIQ TTVVFWPAKLQASSRVVMFRFEFWDCGESA LKKFDHMLLACMENTDAFLFLFSFTDRASFE DLPGQLARIAGEAPGVVRMVIGSKFDQYMHT DVPERDLTAFRQAWELPLLRVKSVPGRRLG 784 2134 A 6308 86 96 GSSPDPASLITMKNQDKKNGAAKQSNPKSSP GQPEAGPEGAQERPSQAAPAVEAEGPGSSQA PRKPEGAQARTAQSGALRDVSEELSRQLEDIL STYCVDNNQGGPGEDGAQGEPAEPEDAEKSR TYVARNGEPEPTPVVNGEKEPSKGDPNTEEIR QSDEVGDRDHRRPQEKKKAKGLGKETILLM QTLNTLSTPEEKLAALCKKYAELLEEHRNSQ	/02	2132	1.	0201	1327	3,73	
KKRKNKSSKRKHRK YSDSDSNSESDTNSDSD DDKKRVKAKKKKKKKKKKKKKKKKKK ESSDSSCKDSEEDLSEATWMEQPNVADTMDL IGPEAPIIHTSQDEKPLKYGHALLPGEGAAMA EYVKAGKRIPRRGEIGLTSEEIGSFECSGYVM SGSRHRRMEAVRLRKENQIYSADEKRALASF NQEERKKESKILASFREMVHKKTKGKDDK 783 2133 A 6305 201 1032 WDDYPQGALRREAAEGLHFLGPPGRVRGQ LRGITGPAWYCHSPSHSLLSAFCHLPTPSRCP AMARPPVPGSVVVPNWHES/RRGQGVPGLHS AQEPPAGVWAA*AASAAALSIDTASYKIFV SGKSGVGKTALVAKLAGLEVPVVHHETTGIQ TTVVFWPAKLQASSRVVMFRFEFWDCGESA LKKFDHMLLACMENTDAFLFLFSPTDRASFE DLPGQLARIAGEAPGVVRMVIGSKFDQYMHT DVPERDLTAFRQAWELPLLRVKSVPGRRLG 784 2134 A 6308 86 96 GSSPDPASLITMKNQDKKNGAAKQSNPKSSP GQPEAGPEGAQERPSQAAPAVEAEGPGSSQA PRKPEGAQARTAQSGALRDVSEELSRQLEDIL STYCVDNNQGGPGEDGAQGEPAEPEDAEKSR TYVARNGEPEPTPVVNGEKEPSKGDPNTEEIR QSDEVGDRDHRRPQEKKKAKGLGKETILLM QTLNTLSTPEEKLAALCKKYAELLEEHRNSQ	İ						
DDKKRVKAKKKKKKKKKKKKKKKKKK ESSDSSCKDSEEDLSEATWMEQPNVADTMDL IGPEAPIIHTSQDEKPLKYGHALLPGEGAAMA EYVKAGKRIPRRGEIGLTSEEIGSFECSGYVM SGSRHRMEAVRLRKENQIYSADEKRALASF NQEERRKRESKILASFREMVHKKTKGKDDK 783 2133 A 6305 201 1032 WDDYPQGALRREAAEGLHFLGPPGRVRGQ LRGITGPAWYCHSPSHSLLSAFCHLPTPSRCP AMARPPVPGSVVVPNWHES/RRGQGVPGLHS AQEPPAGVWAA*AASAAAALISIDTASYKIFV SGKSGVGKTALVAKLAGLEVPVVHHETTGIQ TTVVFWPAKLQASSRVVMFRFEFWDCGESA LKKFDHMLLACMENTDAFLFLFSPTDRASFE DLPGQLARIAGEAPGVVRMVIGSKFDQYMHT DVPERDLTAFRQAWELPLLRVKSVPGRRLG 784 2134 A 6308 86 96 GSSPDPASLITMKNQDKKNGAAKQSNPKSSP GQPEAGPEGAQERPSQAAPAVEAEGPGSSQA PRKPEGAQARTAQSGALRDVSEELSRQLEDIL STYCVDNNQGGPGEDGAQGEPAEPEDAEKSR TYVARNGEPEFIPVVNGEKEPSKGDPNTEEIR QSDEVGDRDHRRPQEKKKAKGLGKEITLLM QTLNTLSTPEEKLAALCKKYAELLEEHRNSQ	1		1	1 1	:		`
ESSDSSCKDSEEDLSEATWMEQPNVADTMDL IGPEAPIIHTSQDEKPLKYGHALLPGEGAAMA EYVKAGKRIPRRGEIGLTSEEIGSFECSGYVM SGSRHRRMEAVRLRKENQIYSADEKRALASF NQEERRKRESKILASFREMVHKKTKGKDDK 783 2133 A 6305 201 1032 WDDYPQGALRREAAEGLHFLGPPGRVRGQ LRGITGPAWYCHSPSHSLLSAFCHLPTPSRCP AMARPPVPGSVVVPNWHES/RRGQGVPGLHS AQEPPAGVWAA*AASAAAALSIDTASYKIFV SGKSGVGKTALVAKLAGLEVPVVHHETTGIQ TTVVFWPAKLQASSRVVMFRFFFWDCGESA LKKFDHMLLACMENTDAFLFFSFTDRASFE DLPGQLARIAGEAPGVVRMVIGSKFDQYMHT DVPERDLTAFRQAWELPLLRVKSVPGRRLG 784 2134 A 6308 86 96 GSSPDPASLITMKNQDKKNGAAKQSNPKSSP GQPEAGPEGAQERPSQAAPAVEAEGPGSSQA PRKPEGAQARTAQSGALRDVSELSRQLEDIL STYCVDNNQGGPGEDGAQGEPAEPEDAEKSR TYVARNGEPEPITPVVNGEKEPSKGDPNTEEIR QSDEVGDRDHRRPQEKKKAKGLGKEITLLM QTLNTLSTPEEKLAALCKKYAELLEEHRNSQ			1				
IGPEAPIIHTSQDEKPLKYGHALLPGEGAAMA EYVKAGKRIPRRGEIGLTSEEIGSFECSGYVM SGSRHRRMEAVRLRKENQIYSADEKRALASF NQEERRKRESKILASFREMVHKKTKGKDDK 783 2133 A 6305 201 1032 WDDYPQGALRREAAEGLHFLGPPGRVRGQ LRGITGPAWYCHSPSHSLLSAFCHLPTPSRCP AMARPPVPGSVVVPNWHES/RRQQGVPGLHS AQEPPAGVWAA*AASAAAALSIDTASYKIFV SGKSGVGKTALVAKLAGLEVPVVHHETTGIQ TTVVFWPAKLQASSRVVMFRFEFWDCGESA LKKFDHMLLACMENTDAFLFLFSFTDRASFE DLPGQLARIAGEAPGVVRMVIGSKFDQYMHT DVPERDLTAFRQAWELPLLRVKSVPGRRLG 784 2134 A 6308 86 96 GSSPDPASLITMKNQDKKNGAAKQSNPKSSP GQPEAGPEGAQERPSQAAPAVEAEGPGSSQA PRKPEGAQARTAQSGALRDVSEELSRQLEDIL STYCVDNNQGGPGEDGAQGEPAEPEDAEKSR TYVARNGEPEPTPVVNGEKEPSKGDPNTEEIR QSDEVGDRDHRRPQEKKKAKGLGKEITLLM QTLNTLSTPEEKLAALCKKYAELLEEHRNSQ	1	i	1		-		
EYVKAGKRIPRRGEIGLTSEEIGSFECSGYVM SGSRHRRMEAVRLRKENQIYSADEKRALASF NQEERKRESKILASFREMVHKKTKGKDDK 783 2133 A 6305 201 1032 WDDYPQGALRREAAEGLHFLGPPGRVRGQ LRGITGPAWYCHSPSHSLLSAFCHLPTPSRCP AMARPPVPGSVVVPNWHES/RRQQGVPGLHS AQEPPAGVWAA*AASAAAALSIDTASYKIFV SGKSGVGKTALVAKLAGLEVPVVHHETTGIQ TTVVFWPAKLQASSRVVMFRFEFWDCGESA LKKFDHMLLACMENTDAFLFLFSFTDRASFE DLPGQLARIAGEAPGVVRMVIGSKFDQYMHT DVPERDLTAFRQAWELPLLRVKSVPGRRLG 784 2134 A 6308 86 96 GSSPDPASLITMKNQDKKNGAAKQSNPKSSP GQPEAGPEGAQERPSQAAPAVEAEGPGSSQA PRKPEGAQARTAQSGALRDVSEELSRQLEDIL STYCVDNNQGGPGEDGAQGEPAEPEDAEKSR TYVARNGEPEPTPVVNGEKEPSKGDPNTEEIR QSDEVGDRDHRRPQEKKKAKGLGKEITLLM QTLNTLSTPEEKLAALCKKYAELLEEHRNSQ					i		
SGSRHRRMEAVRLRKENQIYSADEKRALASF NQEERKKESKILASFREMVHKKTKGKDDK 783 2133 A 6305 201 1032 WDDYPQGALRREAAEGLHFLGPPGRVRGQ LRGITGPAWYCHSPSHSLLSAFCHLPTPSRCP AMARPPVPGSVVVPNWHES/RRGQGVPGLHS AQEPPAGVWAA*AASAAAA\LSIDTASYKIFV SGKSGVGKTALVAKLAGLEVPVVHHETTGIQ TTVVFWPAKLQASSRVVMFRFEFWDCGESA LKKFDHMLLACMENTDAFLFLFSPTDRASFE DLPGQLARIAGEAPGVVRMVIGSKFDQYMHT DVPERDLTAFRQAWELPLLRVKSVPGRRLG 784 2134 A 6308 86 96 GSSPDPASLITMKNQDKKNGAAKQSNPKSSP GQPEAGPEGAQERPSQAAPAVEAEGPGSSQA PRKPEGAQARTAQSGALRDVSEELSRQLEDIL STYCVDNNQGGPGEDGAQGEPAEPEDAEKSR TYVARNGEPEPTPVVNGEKEPSKGDPNTEEIR QSDEVGDRDHRRPQEKKKAKGLGKEITLLM QTLNTLSTPEEKLAALCKKYAELLEEHRNSQ	1	1	1		l	l	
783 2133 A 6305 201 1032 WDDYPQGALRREAAEGLHFLGPPGRVRGQ LRGITGPAWYCHSPSHSLLSAFCHLPTPSRCP AMARPPVPGSVVVPNWHES/RRGQGVPGLHS AQEPPAGVWAA*AASAAAA/LSIDTASYKIFV SGKSGVGKTALVAKLAGLEVPVVHHETTGIQ TTVVFWPAKLQASSRVVMFRFFFWDCGESA LKKFDHMLLACMENTDAFLFLFSFTDRASFE DLPGQLARIAGEAPGVVRMVIGSKFDQYMHT DVPERDLTAFRQAWELPLLRVKSVPGRRLG 784 2134 A 6308 86 96 GSSPDPASLITMKNQDKKNGAAKQSNPKSSP GQPEAGPEGAQERPSQAAPAVEAEGPGSSQA PRKPEGAQARTAQSGALRDVSELLSRQLEDIL STYCVDNNQGGPGEDGAQGEPAEPEDAEKSR TYVARNGEPEPTPVVNGEKEPSKGDPNTEEIR QSDEVGDRDHRRPQEKKKAKGLGKEITLLM QTLNTLSTPEEKLAALCKKYAELLEEHRNSQ			1				
783 2133 A 6305 201 1032 WDDYPQGALRREAAEGLHFLGPPGRVRGQ LRGITGPAWYCHSPSHSLLSAFCHLPTPSRCP AMARPPVPGSVVVPNWHES/RRGQGVPGLHS AQEPPAGVWAA*AASAAAA\LSIDTASYKIFV SGKSGVGKTALVAKLAGLEVPVVHHETTGIQ TTVVFWPAKLQASSRVVMFRFEFWDCGESA LKKFDHMLACMENTDAFLFLFSFTDRASFE DLPGQLARIAGEAPGVVRMVIGSKFDQYMHT DVPERDLTAFRQAWELPLLRVKSVPGRRLG GSPDPASLITMKNQDKKNGAAKQSNPKSSP GQPEAGPEGAQERPSQAAPAVEAEGPGSSQA PRKPEGAQARTAQSGALRDVSEELSRQLEDIL STYCVDNNQGGPGEDGAQGEPAEPEDAEKSR TYVARNGEPEPTPVVNGEKEPSKGDPNTEEIR QSDEVGDRDHRRPQEKKKAKGLGKEITLLM QTLNTLSTPEEKLAALCKKYAELLEEHRNSQ	1						7
LRGITGPAWYCHSPSHSLLSAFCHLPTPSRCP AMARPPVPGSVVVPNWHES/RRGQGVPGLHS AQEPPAGVWAA*AASAAAA\LSIDTASYKIFV SGKSGVGKTALVAKLAGLEVPVVHHETTGIQ TTVVFWPAKLQASSRVVMFRFEFWDCGESA LKKFDHMLLACMENTDAFLFLFSFTDRASFE DLPGQLARIAGEAPGVVRMVIGSKFDQYMHT DVPERDLTAFRQAWELPLLRVKSVPGRRLG 784 2134 A 6308 86 96 GSSPDPASLITMKNQDKKNGAKQSNPKSSP GQPEAGPEGAQERPSQAAPAVEAEGPGSSQA PRKPEGAQARTAQSGALRDVSEELSRQLEDIL STYCVDNNQGGPGEDGAQGEPAEPEDAEKSR TYVARNGEPEPTPVVNGEKEPSKGDPNTEEIR QSDEVGDRDHRRPQEKKKAKGLGKEITLLM QTLNTLSTPEEKLAALCKKYAELLEEHRNSQ	<u> </u>		ــــــــــــــــــــــــــــــــــــــ				
AMARPPVPGSVVVPNWHES/RRGQGVPGLHS AQEPPAGVWAA*AASAAAA\LSIDTASYKIFV SGKSGVGKTALVAKLAGLEVPVVHHETTGIQ TTVVFWPAKLQASSRVVMFRFEFWDCGESA LKKFDHMLLACMENTDAFLFLFSFTDRASFE DLPGQLARIAGEAPGVVRMVIGSKFDQYMHT DVPERDLTAFRQAWELPLLRVKSVPGRRLG 784 2134 A 6308 86 96 GSSPDPASLITMKNQDKKNGAAKQSNPKSSP GQPEAGPEGAQERPSQAAPAVEAEGPGSSQA PRKPEGAQARTAQSGALRDVSEELSRQLEDIL STYCVDNNQGGPGEDGAQGEPAEPEDAEKSR TYVARNGEPEPTPVVNGEKEPSKGDPNTEEIR QSDEVGDRDHRRPQEKKKAKGLGKEITLLM QTLNTLSTPEEKLAALCKKYAELLEEHRNSQ	783	2133	Į A	6305	201	1032	
AQEPPAGVWAA*AASAAAA\LSIDTASYKIFV SGKSGVGKTALVAKLAGLEVPVVHHETTGIQ TTVVFWPAKLQASSRVVMFRFEFWDCGESA LKKPDHMLLACMENTDAFLFLFSFTDRASFE DLPGQLARIAGEAPGVVRMVIGSKFDQYMHT DVPERDLTAFRQAWELPLLRVKSVPGRRLG 784 2134 A 6308 86 96 GSSPDPASLITMKNQDKKNGAAKQSPPKSSP GQPEAGPEGAQERPSQAAPAVEAEGPGSSQA PRKPEGAQARTAQSGALRDVSEELSRQLEDIL STYCVDNNQGGPGEDGAQGEPAEPEDAEKSR TYVARNGEPEPTPVVNGEKEPSKGDPNTEEIR QSDEVGDRDHRRPQEKKKAKGLGKEITLLM QTLNTLSTPEEKLAALCKKYAELLEEHRNSQ	1	1	1			į	
AQEPPAGVWAA*AASAAAA\LSIDTASYKIFV SGKSGVGKTALVAKLAGLEVPVVHHETTGIQ TTVVFWPAKLQASSRVVMFRFEFWDCGESA LKKPDHMLLACMENTDAFLFLFSFTDRASFE DLPGQLARIAGEAPGVVRMVIGSKFDQYMHT DVPERDLTAFRQAWELPLLRVKSVPGRRLG 784 2134 A 6308 86 96 GSSPDPASLITMKNQDKKNGAAKQSPPKSSP GQPEAGPEGAQERPSQAAPAVEAEGPGSSQA PRKPEGAQARTAQSGALRDVSEELSRQLEDIL STYCVDNNQGGPGEDGAQGEPAEPEDAEKSR TYVARNGEPEPTPVVNGEKEPSKGDPNTEEIR QSDEVGDRDHRRPQEKKKAKGLGKEITLLM QTLNTLSTPEEKLAALCKKYAELLEEHRNSQ							
TTVVFWPAKLQASSRVVMFRFEFWDCGESA LKKFDHMLLACMENTDAFLFLFSFTDRASFE DLPGQLARIAGEAPGVVRMVIGSKFDQYMHT DVPERDLTAFRQAWELPLLRVKSVPGRRLG 784 2134 A 6308 86 96 GSSPDPASLITMKNQDKKNGAAKQSNPKSSP GQPEAGPEGAQERPSQAAPAVEAEGPGSSQA PRKPEGAQARTAQSGALRDVSEELSRQLEDIL STYCVDNNQGGPGEDGAQGEPAEPEDAEKSR TYVARNGEPEPTPVVNGEKEPSKGDPNTEEIR QSDEVGDRDHRRPQEKKKAKGLGKEITLLM QTLNTLSTPEEKLAALCKKYAELLEEHRNSQ	1						
TTVVFWPAKLQASSRVVMFRFEFWDCGESA LKKFDHMLLACMENTDAFLFLFSFTDRASFE DLPGQLARIAGEAPGVVRMVIGSKFDQYMHT DVPERDLTAFRQAWELPLLRVKSVPGRRLG 784 2134 A 6308 86 96 GSSPDPASLITMKNQDKKNGAAKQSNPKSSP GQPEAGPEGAQERPSQAAPAVEAEGPGSSQA PRKPEGAQARTAQSGALRDVSEELSRQLEDIL STYCVDNNQGGPGEDGAQGEPAEPEDAEKSR TYVARNGEPEPTPVVNGEKEPSKGDPNTEEIR QSDEVGDRDHRRPQEKKKAKGLGKEITLLM QTLNTLSTPEEKLAALCKKYAELLEEHRNSQ	ŀ						SGKSGVGKTALVAKLAGLEVPVVHHETTGIQ
DLPGQLARIAGEAPGVVRMVIGSKFDQYMHT DVPERDLTAFRQAWELPLLRVKSVPGRRLG 784 2134 A 6308 86 96 GSSPDPASLITMKNQDKKNGAAKQSNPKSSP GQPEAGPEGAQERPSQAAPAVEAEGPGSSQA PRKPEGAQARTAQSGALRDVSEELSRQLEDIL STYCVDNNQGGPGEDGAQGEPAEPEDAEKSR TYVARNGEPEPTPVVNGEKEPSKGDPNTEEIR QSDEVGDRDHRRPQEKKKAKGLGKEITLLM QTLNTLSTPEEKLAALCKKYAELLEEHRNSQ							TTVVFWPAKLQASSRVVMFRFEFWDCGESA
DLPGQLARIAGEAPGVVRMVIGSKFDQYMHT DVPERDLTAFRQAWELPLLRVKSVPGRRLG 784 2134 A 6308 86 96 GSSPDPASLITMKNQDKKNGAAKQSNPKSSP GQPEAGPEGAQERPSQAAPAVEAEGPGSSQA PRKPEGAQARTAQSGALRDVSEELSRQLEDIL STYCVDNNQGGPGEDGAQGEPAEPEDAEKSR TYVARNGEPEPTPVVNGEKEPSKGDPNTEEIR QSDEVGDRDHRRPQEKKKAKGLGKEITLLM QTLNTLSTPEEKLAALCKKYAELLEEHRNSQ	-	1	1		ļ	!	
784 2134 A 6308 86 96 GSSPDPASLITMKNQDKKNGAAKQSNPKSSP GQPEAGPEGAQERPSQAAPAVEAEGPGSSQA PRKPEGAQARTAQSGALRDVSEELSRQLEDIL STYCVDNNQGGPGEDGAQGEPAEPEDAEKSR TYVARNGEPEPTPVVNGEKEPSKGDPNTEEIR QSDEVGDRDHRRPQEKKKAKGLGKEITLLM QTLNTLSTPEEKLAALCKKYAELLEEHRNSQ		1		1			
784 2134 A 6308 86 96 GSSPDPASLITMKNQDKKNGAAKQSNPKSSP GQPEAGPEGAQERPSQAAPAVEAEGPGSSQA PRKPEGAQARTAQSGALRDVSEELSRQLEDIL STYCVDNNQGGPGEDGAQGEPAEPEDAEKSR TYVARNGEPEPTPVVNGEKEPSKGDPNTEEIR QSDEVGDRDHRRPQEKKKAKGLGKEITLLM QTLNTLSTPEEKLAALCKKYAELLEEHRNSQ							
GQPEAGPEGAQERPSQAAPAVEAEGPGSSQA PRKPEGAQARTAQSGALRDVSEELSRQLEDIL STYCVDNNQGGPGEDGAQGEPAEPEDAEKSR TYVARNGEPEPTPVVNGEKEPSKGDPNTEEIR QSDEVGDRDHRRPQEKKKAKGLGKEITLLM QTLNTLSTPEEKLAALCKKYAELLEEHRNSQ	784	2134	A	6308	86	96	
PRKPEGAQARTAQSGALRDVSEELSRQLEDIL STYCVDNNQGGPGEDGAQGEPAEPEDAEKSR TYVARNGEPEPTPVVNGEKEPSKGDPNTEEIR QSDEVGDRDHRRPQEKKKAKGLGKEITLLM QTLNTLSTPEEKLAALCKKYAELLEEHRNSQ] ',	-13-	1	55.00			
STYCVDNNQGGPGEDGAQGEPAEPEDAEKSR TYVARNGEPEPTPVVNGEKEPSKGDPNTEEIR QSDEVGDRDHRRPQEKKKAKGLGKEITLLM QTLNTLSTPEEKLAALCKKYAELLEEHRNSQ		[1	[[ĺ	
TYVARNGEPEPTPVVNGEKEPSKGDPNTEEIR QSDEVGDRDHRRPQEKKKAKGLGKEITLLM QTLNTLSTPEEKLAALCKKYAELLEEHRNSQ			1	1			
QSDEVGDRDHRRPQEKKKAKGLGKEITLLM QTLNTLSTPEEKLAALCKKYAELLEEHRNSQ			-				
QTLNTLSTPEEKLAALCKKYAELLEEHRNSQ		1					
						1	
L KQMKLLQKKQSQLVQEKDHLKGEHSKAVLA	1		1]]			,
	L			لـــــا			KYMKLLYKKYSYLVYEKDHLRGEHSKAVLA

000 10	LOBOTE	т	Tone			T
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	J	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	1		ļ	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
F		l	i	peptide	Soquano	/=possible nucleotide deletion, \=possible
	J	ļ	J			
— —	 	 		sequence		nucleotide insertion
			1			RSKLESLCRELQRHNRSLKEEGVQRAREEEE
		1				KRKEVTSHFQVTLNDIQLQMEQHNERNSKLR
	į.	j	ļ	J		QENMELAERLKKLIEQYELREEHIDKVFKHK
1		İ				DLQQQLVDAKLQQAQEMLKEAEERHOREKD
!		l				FLLKEAVESQRMCELMKQQETHLKQQLALY
1			l			TEKFEEFONTLSKSSEVFTTFKOEMEKMTKKI
Ì		ł	1			KKLEKETTMYRSRWESSNKALLEMAEEKTV
		ļ			•	RDKELEGLQVKIQRLEKLCRALQT/GAQ*PVR
						GQRWGSHRTSAVRIFS
785	2135	A	6319	1402	000	
/ 63	2133	Α.	0319	1493	889	SPQGPLLRSVSPVSAGASSVTPGGAQPGVTTT
}		}			ļ	PPSLVAVAPAPGSAAGPAAGWQ+HAGCR/WT
		-				KLPWSWGMRPMKIFFSEEYRSISTRISHDAL*
						EKCTQPAKPLSMIR\TGSSVSPG/PLVKWNWT
ļ		1				RREFRNSGTRVVSSCCGMSCMYSFLGHCSV/S
1	1	ł				QDLPLVHVDVGWQPPLGPTVGLRPGLLPLHD
1	1		1			TTPCQKLVVDDLDWA
786	2136	A	6320	551	135	RWLPVAECDSSCVGCTGEGPGNCKECISGYA
1 - 3			5525	331	155	REHGQCADVDECSLAEKTCVRKNENCYNTP
		-				GSYVCVCPDGFEET/RRCLCAAGRG*SHRRRK
		ļ	1			
į.			1			PDTAALPRRPVMCRTYPLNYSEGCPVENVAL
		<u> </u>				RMPSPAVDSGGERLPAL
787	2137	A	6330	1693	227	DYVLTAELHRQRSPGVSFGLSVFNLMNAIMG
ł	1.	i	ł .			SGILGLAYVMANTGVFGFSFLLLTVALLASYS
	1					VHLLLSMCIQTAYLGP*TNYFMVLPAH*LTCL
						PLIEFLQSL*NSL*AVTSYEDLGLFAFGLPGKL
		1				VVAGTIIIQNIGAMSSYLLIIKTELPAAIAEFLT
i			1			GDYSRYWYLDGQTLLIIICVGIVFPLALLPKIG
1	1	ŀ	1			FLGYTSSLSFFFMMFFALVVIIKKWSIPCPLTL
						NYVEKGFQISNVTDDCKPKLFHFSKESAYALP
	ł					
	1					TMAFSFLCHTSILPIYCELQSPSKKRMQNVTN
						TAIALSFLIYFISALFGYLTFYD/GTTKAQRGE
i						VTCHRIKDKVESELLKG***IP*SHDVVVMT\V
		1		i		KLCILFAVLL\TVPLIHFPARKAVTMMFFSNFP
						FSWIRHFLITLALNIIIVLLAIYVPDIRNVFGVV
		[[GASTSTCLIFIFPGLFYLKLSREDFLSWKKLGV
						GCFC/LLSFKTSILRNSLSVYIILPASRKSIYFKI
788	2138	A	6351	1	6622	PRSLCFSLWAEAAVLADGGLRRRRRLLRGTM
ļ						SASFVPNGASLEDCHCNLFCLADLTGIKWKK
1		1			<u> </u>	YVWQGPTSAPILFPVTEEDPILSSFSRCLKADV
1	İ	1			ì	LG/VWRRDQRPERRE\L*IFWGGEDP\VLLTLF
1		1			·	TMTYQKKKMECGRMDFPMNAVLCFSKAVH
						MI I EDOLAGDARDIOUNDICUMAVECTSKAVH
	ļ]]			NLLERCLMNRNFVRIGKWFVKPYEKDEKPIN
	1	ļ.				KSEHLSCSFTFFLHGDSNVCTSVEINQHQPVY
ł					1	LLSEEHITLAQQSNSPFQVILCPFGLNGTLTGQ
		1				AFKMSDSATKKLIGEWKQFYPISCCLKEMSE
i						EKQEDMDWEDDSLAAVEVLVAGVRMIYPAC
						FVLVPQSDIPTPSPVGSTHCSSSCLGVHQVPAS
1	1	l				TRDPAMSSVTLTPPTSPEEVQTVDPQSVQKW
		l				VKFSSVSDGFNSDSTSHHGGKIPRKLANHVV
	1	1		•		DRVWQECNMNRAQNKRKYSASSGGLCEEAT
		1				
1		[[ĺ	1	AAKVASWDFVEATQRTNCSCLRHKNLKSRN
l l					1	AGQQGQAPSLGQQQQILPKHKTNEKQEKSEK
1 '				l	ļ	PQKRPLTPFHHRVSVSDDVGMD\ADS\ASQRL
1				l	1	V\ISAP\DSQ\VRFSNIR\TNDVAK\TPQMHGTE
1				l		MANSPQPPPLSP\HPCDVVDEGVTKTPSTPQS
ł				ł	ŀ	QHFYQMPTPDPLVPSKPMEDRIDSLSQSFPPO
	1					YQEAVEPTVYVGTAVNLEEDEANIAWKYYK
						FPKKKDVEFLPPQLPSDKFKDDPVGPFGOESV
				l		TSVTELMVQCKKPLKVSDELVQQYQIKNQCL
1				ł	ì	SAIASDAEQEPKIDPYAFVEGDEEFLFPDKKD
	J	L				OUTUODATE OF LATE LATE AT A SECURE LITER DEVIL

CODO TO	OPC TO) A dist	1000	D., 4: • •	T 5	1
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-]	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ļ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ł		peptide		/=possible nucleotide deletion, \=possible
	 			sequence	ļ	nucleotide insertion
						RONSEREAGKKHKVEDGTSSVTVLSHEEDA
		1				MSLFSPSIKQDAPRPTSHARPPSTSLIYDSDLA
ĺ	i	1		1	İ	VSYTDLDNLFNSDEDELTPGSKRSANGSDDK
i						ASCKESKTGNLDPLSCISTADLHKMYPTPPSL
						EQHIMGFSPMNMNNKEYGSMDTTPGGTVLE
						GNSSSIGAQFKIEVDEGFCSPKPSEIKDFSYVY
						KPENCQILVGCSMFAPLKTLPSQYLPLIKLPEE
ļ		1		1		CIYRQSWTVGKLELLSSGPSMPFIKEGDGSNM
		ĺ				DQEYGTAYTPQTHTSCGMPPSSAPPSNSGAGI
	1	ļ		i		LPSPSTPRFPTPRTPRTPRTPRGAGGPASAQGS
	ļ			1	ļ	VKYENSDLYSPASTPSTCRPLNSVEPATVPSIP
				•		EAHSLYVNLILSESVMNLFKDCNSDSCCICVC
		ŀ	Ì			NMNIKGADVGVYIPDPTQEAQYRCTCGFSAV
		}				MNRKFGNNSGLFFEDELDIIGRNTDCGKEAE
	ļ	Ì				KRFEALRATSAEHVNGGLKESEKLSDDLILLL
	ĺ					QDQCTNLFSPFGAADQDPFPKSGVISNWVRV
İ						EERDCCNDCYLALEHGRQFMDNMSGGKVDE
[ĺ		((ALVKSSCLHPWSKRNDVSMQCSQDILRMLLS
		1	}			LQPVLQDAIQKKRTVRPWGVQGPLTWQQFH
						KMAGRGSYGTDESPEPLPIPTFLLGYDYDYLV
						LSPFALPYWERLMLEPYGSQRDIAYVVLCPE
		İ		,		NEALLNGAKSFFRDLTAIYESCRLGQHRPVSR
		1				LLTDGIMRVGSTASKKLSEKLVAEWFSQAAD
		[ŀ		GNNEAFSKLKLYAQVCRYDLGPYLASLPLDS
		}				SLLSQPNLVAPTSQSLITPPQMTNTGNANTPS
		ŀ				ATLASAASSTMTVTSGVAISTSVATANSTLTT
						ASTSSSSSNLNSGVSSNKLPSFPFFGSMNSNA
1						AGSMSTQANTVQSGQLGGQQTSALQTAGISG
				-		ESSSLPTQPHPDVSESTMDRDKVGIPTDGDSH
						AVTYPPAIVVYIIDPFTYENTDESTNSSSVWTL
						GLLRCFLEMVQTLPPHIKSTVSVQIIPCQYLLQ PVKHEDREIYPQHLKSLAFSAFTQCRRPLPTS
			1			TNVKTLTGFGPGLAMETALRSPDRPECIRLYA
						PPFILAPVKDKQTELGETFGEAGQKYNVLFV
						GYCLSHDQRWILASCTDLYGELLETCIINIDVP
						NRARRKKSSARKFGLQKLWEWCLGLVOMSS
						LPWRVVIGRLGRIGHGELKDWSCLLSRRNLQ
						SLSKRLKDMCRMCGISAADSPSILSACLVAM
						EPQGSFVIMPDSVSTGSVFGRSTTLNMQTSQL
						NTPQDTSCTHILVFPTSASVQVASATYTTENL
	,				1	DLAFNPNNDGADGMGIFDLLDTGDDLDPDII
						NILPASPTGSPVHSPGSHYPHGGDAGKGOSTD
						RLLSTEPHEEVPNILQQPLALGYFVSTAKAGP
						LPDWFWSACPQAQYQCPLFLKASLHLHVPSV
						QSDELLHSKHSHPLDSNQTSDVLRFVLEQYN
				i		ALSWLTCDPATQDRRSCLPIHFVVLNQLYNFI
i						MNML
789	2139	A	6359	1	2002	TGTLTEDGLDVMGVVPLKGQAFLPLVPEPRR
				*	2002	LPVGPLLRALATCHALSRLQDTPVGDPMDLK
						MVESTGWVLEEEPAADSAFGTQVLAVMRPP
-						LWEPOLOAMEEPPVPVSVLHRFPFSSALORM
				j	1	SVVVAWPGATQPEAYVKGSPELVAGLCNPET
1				ľ		VPTDFAQMLQSYTAAGYRVVALASKPLPSVP
ļ]]	l		SLEAAQQLTRDTVEGDLSLLGLLVMRNLLKP
	1			ĺ		QTTPVIQALRRTRIRAVMVTGDNLQTAVTVA
	ļ			1		, ,
						RGCGMVAPQEHLIIVHATHPERGQPASLEFLP
ļ						MESPTAVNGVKDPDQAASYTVEPDPRSRHLA
	1	İ		_	}	LSGPTFGIIVKHFPKLLPKVLVQGTVFARMAP
1	Į			-	1	EQKTELVCELQKLQYCVGMCGDGANDCGAL
			L			KAADVGISLSQAEASVVSPFTSSMASIECVPM

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion VIREGRCSLDTSFSVFKYMALYSLTQFISVLIL
						YTINTNLGDLQFLAIDLVITTTVAVLMSRTGP ALVLGRVRPPGALLSVPVLSSLLLQMVLVTG VQLGGYFLTLAQPWFVPLNRTVAAPDNLPNY ENTVVFSLSSFQYLILAAAVSKGAPFR\RPLTN NVPFLLASAL*SSVLVVLVLSPGLLHGPLALR NITDTGFKLLLVGLVTLNFVGGLHAGERARP VPPRLPAPPPAQAG\SKKRFKQLERELAEQPW PPLPAGPLR
790	2140	A	6380	76	1059	SSAGSARKLQVMALAARLWRLLPFRRGAAP GSRLPAGTSGSRGHCGPCRFRGFEVMGNPGT FKRGLLLSALSYLGFETYQVISQAAVVHATA KVEEILEQADYLYESGETEKLYQLLTQYKESE DAELLWRLARASRDVAQLSRTSEEEKKLLVY EALEYAKRA/L/EKNESSFASHKWYAICLSDV GDYEGIKAKIANAYIIKEHFEKAIELNPKDATS IHLMGIWCYTFAEMPWYQRRIA*NACLQLPP *FPPYEKALG\YFHRAEQVDPNFYSKNLLLLG KTYLKLHNKKLAAFWLMKAKDYPAHTEED KQIQTEAAQLLTSFSEKN
791	2141		6434	3	1460	IALLIVDGLAWDDQGGLALLHISPSKLIL*QDS SGMS/YVMVRCTITRAFFKSLLCHICQYSIGPQ *VT\CPGQDACKE*KSTAN*GG*RE**PQVLFF AFLSNPAVKFGRMSKKQRDSLYAEVQKHQQ RLQEQRQQQSGEAEALARVYSSSISNGLSNLN NETSGTYANGSVIDLPKSEGYYNVVSGQPSP DQSGLDMT\GIKQIKQEPIYDLTSVPNLFTY\SS FNN\GQLAPGIT\MTEIDRIAQNIIKSHLETCQY TMEELHQLAWQTHTYEEIKAYQSKSREALW QQCAIQITHAIQYVVEFAKRITGFMELCQNDQ ILLLKSGCLEVVLVRMCRAFNPLNNTVLFEO
					-	KYGGMQMFKALGSDDLVNEAFDFAKNLCSL QLTEEEIALFSSAVLISPDRAWLIEPRKVQKLQ EKIYFALQHVIQKNHLDDETLAKLIAKIPTITA VCNLHGEKLQVFKQSHPEIVNTLFPPLYKELF NPDCATACK
792	2142	Α	6440	92	781	SRGTFRCFCRDFFPCFSNMRLFLWNAVLTLFV TSLIGALIPEPEVKIEVLQKPFICHRKTKGGDL MLVHYEGYLEKDGSLFHSTHKHNNGQPIWFT LGILEALKGWGPGA*K/DMCVGEKRKLIIPPA LGYGKEGKGKIPPESTLIFNIDLLEIRNGPRSH ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH GAVVNESHHDALVEDIFDKEDEDKDGFISAR EFTYKHDEL
793	2143	Α	6446		152	PRLKRI.VVTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEKWSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITC\LEWDQSGFPGFLFSRWPTGQIK\CWS MGVSTLA\NSWE\SSVGSL\VEGGPHLWALS\ WLH\NGVKLALHVEKSGASSFGEKFSR\VKFS P\SLTLF\GGNAMEGWIAVTVSGLVTVSLLQ\P SGQVL\TST\ESLCRLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRLSATNDLDRVSA VALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD EPAMKRPRTAGPAVHLKAMQLSWTSLALVG

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide deletion, \=possible nucleotide insertion IDSHGKLSV\LRLSPSMGHPLEVGLALRHLLFL LEYCMVTGYDWWDILLHVQPSMVQSLVEKL HEEYTRQTAALQQVLSTRILAMKASLCKLSP CTVTRVCDYHTKLFLIAISSTLKSLLRPHFLNT PDKSPGDRLTEICTKITDVDIDKVMINLKTEEF VLDMNTLQALQQLLQWVGDFVLYLLASLPN QPCPTSEPCPTSEPSPTSEPSPTSEPSSP*SLCG SLLRPGHSFLRDGTSLGMLRELMVVIRIWGLL KPSCLPVYTATSDTQDSMSLLFRLLTKLWICC RDEGPASEPDEALVDECCLLPSQLLIPSLDWL PASDGLVSRLQPKQPLRLQFGRAPTLPGSAAT LQLDGLARAPGQPKIDHLRRLHLGACPTEEC KACTRCGCVTMLKSPNRTTAVKQWEQRWIK NC/LVRWALVAGAPQLPLSPAAPQLLLSYPSA APEPGCCKSHRSPWTLLGAVNLSPPCRAVEG
794	2144	A	6490	418	585	PRSLDHLHPEDRP NGDKADLENESCRAQVLMPVVPALWEAEGG
795	2145	A	6499	395	1027	GSIEPRDLRLQ*AVITPL\TPAWVTQ KLLWLPPHSEQKRSPLYHPQGPSGTTPSAP\FS
						SHSPPPSLLQA\PSIAAFLRTHGHISASGPLRMP FPH/H*NAFLLVFPGQRSQLTS/PSHYLCREVFP DHHHHLCRLSLESSPLFHHRVLFCVPKQNVN STRAQIFCLFVHIVGCRCINTFPLHLFRLHLWL HFLQIPLCKKNKSVKLGKTVVGRGCQSAAGS DTRVRAAVGAPGLPVEPLV
796	2146	Α	6503	68	936	HSALLTHSSFCVFTLCQDFFTYSSMSEEVTYA DLQFQNSSEMEKIPEIGKFGEKAPPAPSHVWR PAALFLTLLCLLLIGLGVLASMFHVTLKIEM KKMNKLQNISEELQRNISLQLMSNMNISNKIR NLSTTLQTIATKLCRELYSKEQEHKCKPCPRR WIWHKDSCYFLSDDVQTWQESKMACAAQN ASLLKINNKNALEFIKSQSRSYDYWLGLSPEE DSYSWYESG*YNQVPSAWVIRNAPDLNNMY CGYINRLYVQYYHCTYKQRMICEKMANPVQ LGSTYFREA
	2147	A .	6507	1	881	PGSTHASARSQVPRSAGEAAPHSRRPPGLLPH APRAASAQLEERMRDPHPGMTLQEGDCRGS QTVSLTMGTADSDEMAPEAPQHTHIDVHIHQ ESALAKLLLTCCSALRPRATQARGSSRLLVAS WVMQIVLGILSAVLGGFFYIRDYTLLVTSGA AIWTGAVAVLAGAAAFIYEKRGGTYWALLR TLLALAAFSTAIAALKLWNEDFRYGYSYYNS ACRISSSSDWNTPAPTQSPEEVRRLHLCTSFM DMLKALFRTLQAMLLGVWILLLLASLTPLWL /SL/RGECSQPKG*VPKKRDQKEMLEVSGI*PG STHASARSQVPRSAGEAAPHSRRPPGLLPHAP RAASAQLEERMRDPHPGMTLQEGDCRGSQT VSLTMGTADSDEMAPEAPQHTHIDVHIHQES ALAKLLLTCCSALRPRATQARGSSRLLVASW VMQIVLGILSAVLGGFFYIRDYTLLVTSGAAI WTGAVAVLAGAAAFIYEKRGGTYWALLRTL LALAAFSTAIAALKLWNEDFRYGYSYYNSAC RISSSSDWNTPAPTQSPEEVRRLHLCTSFMDM LKALFRTLQAMLLGVWILLLASLTPLWLYC
798	2148	A	6528	912	2287	WRMFPTKGVSP VPNYLPSVSSAIGGEVPQRYVWRFCIGLHSAP RFLVAFAYWNHYLSCTSPCSCYRPLCRLNFG LNVVENLALLVLTYVSSSEDF/TWVPG*GRSG

SEQ ID	SEO ID	- Nat	CEO	<u> </u>	Le company	
NO: of	NO: of	Met hod	SEQ ID NO:	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	1100	in in	beginning nucleotide	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
eotide	seq-		USSN	location	location	F=Phenylalanine, G=Glycine, H=Histidine,
seq-	uence		09/496		corresponding	I=Isoleucine, K=Lysine, L=Leucine,
uence	acnoe		914	correspondi	to last amino acid residue	M=Methionine, N=Asparagine, P=Proline,
defice		}	714	ng to first amino acid		Q=Glutamine, R=Arginine, S=Serine,
1 1			1		of peptide	T=Threonine, V=Valine, W=Tryptophan,
1 1				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
i l				peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
						EVFPEGTGLPLPHSDLPTSWCGHSLQCGSQSS
						FPPAIHENAFIVFIASSLGHMLLTCILWRLTKK
						HTVSQE\DGLSLAGAPRQPRRKSRTSVLRIRV
					ļ	MVRWELSSNGNPGRGVLGLGLGLGNKLRVV
					ļ	GQNLGL*HCVWVVWETGE*KRWRLQMGIE*
						GVASRRQ*VRNSVRGLVCHNSSAPPMYMGFF
]						SPTVFGGGVGG*LHVTFILHPPEVEAAGIPLLL
					1	GPSLPQRQGREHIVVILAAPACAPFHDR*WEP
						REIRPSP*ELGLRGEPTLSYPASCRVIRQPIP*D
						RKSYSWKQRLFIINFISFFSALAVYFRHNMYC
						EAGVYTIFAILEYTVVLTNMAFHMTAWWDF
700	0140					GNKELLITSQPEEKRF
799	2149	Α	6529	1	874	FFFFQRINFIEHSGSVSLLALACDLGWCEDWS
						CCLVQGGGDLVDVVQTNHGEDEAGGDTDSV
						DEARCKESQQEAQENLREDLCLESFAKDKIL
						QIIEGSEREHEETRTKQAALDGEPLGGGQLTA
						VHLHPSKEQQGQEGGERQRGARTHHWRGW
				:		EKGRRVRLRPPSGKLRADQPVRKLGGPTPS/T
1						ELPGLQPHAPTPHTA/PATPTYSPAPDTPNPPV
1						RWKCPLPVEPRTRQLCRERTRKACPPKPRPPL
l		İ				GLPGDPTGPVTHHAPPVSPTGASGQERRAEP
				_		GAVSYAHASATK
800	2150	Α	6544	2	662	SAQRWAAVAGRWGCRLLALLLLVPGPGGAS
						EITFELPDNAKQCFYEDIAQGTKCTLEFQVITG
						GHYDVDCRLEDPDGKVLYKEMKKQYDSFTF
						TASKNGTYKFCFSNE\FSTFTHKTVYFDFQVG
- 1		-				E\THLCFLVR/DRVSALTQMESACVSIHEALKS
		i				VIDYQTHFRLREAQGRSRAEDLNTRVAYWSV
				1		GEALILLVVSIGQVFLLKSFFSDKRTTTTRVGS
108	2151	Α	6556	1	1319	TPCMECIKGEGLREPQNLSGSQREPQTEGSM
						DGWRRMPRWGLLLLLWGSCTFGLPTDTTTF
i	1	- 1				KRIFLKRMPSIRESLKERGVDMARLGPEWSOP
						MKRLTLGNTTSSVILTNYMDTQYYGEIGIGTP
						PQTFKVVFDTGSSNVWVPSSKCSRLYTACVY
	1	.	1			HKLFDASDSSSYKHNGTELTLRYSTGTVSGFL
,	1			ļ		SQDIITVGGITVTQMFGEVTEMPALPFMLAEF
	į		ļ	ļ	•	DGVVGMGFIEQAIGRVTPIFDNIISQGVLKED
	ŀ	- 1				VFSFYYNRDSENSQSLGGQIVLGGSDPOHYE
			l	ļ		GNFHYINLIKTGVWQIQMKGVSVGSSTLLCE
		1	1			DGCLALVDTGASYISGSTSSIEKLMEALGAKE
	ļ					KRLFDYVVKCNEGPTLPPTFLFLLGGKDTPLT
	i					SADYLFQESYSSKKLSTLAIHAMYIPPPTGPTL
		1		İ		\ALGATF\IRKFYTEFDRGNNPHGFALAR
802	2152	A	6567	13	6147	MCLGRMGASSPRSPEPVGPPAPGLPFCCGGSL
. 1	- J	-				LAVVVLLALPVAWGQCNAPEWLPFARPTNL
- 1					ĺ	TDEFEFPIGTYLNYECRPGYSGRPFSIICLKNS
1		İ				VWTGAKDRCRRKSCRNPPDPVNGMVHVIKG
1		[1		IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW
ļ					ĺ	DNETPICDRIPCGLPPTITNGDFISTNRENFHY
1						GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND
j	ļ		ļ	1		
- 1		ļ		1		DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD
1		[l		NRSLFSLNEVVEFRCQPGFVMKGPRRVKCQA
1	- 1	1		1		LNKWEPELPSCSRVCQPPPDVLHAERTQRDK
		- 1				DNFSPGQEVFYSCEPGYDLRGAASMRCTPQG
				l	}	DWSPAAPTCEVKSCDDFMGQLLNGRVLFPV
				l l		NLQLGAKVDFVCDEGFQLKGSSASYCVLAG
			i		I	
-						MESLWNSSVPVCEQIFCPSPPVIPNGRHTGKP
						MESLWNSSVPVCEQIFCPSPPVIPNGRHTGKP LEVFPFGKAVNYTCDPHPDRGTSFDLIGESTIR
						MESLWNSSVPVCEQIFCPSPPVIPNGRHTGKP

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	Į.	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	ļ			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
-				peptide		/=possible nucleotide deletion, \=possible
			<u> </u>	sequence		nucleotide insertion
						SITCLDNLVWSSPKDVCKRKSCKTPPDPVNG
1						MVHVITDIQVGSRINYSCTTGHRLIGHSSAECI
	ŀ					LSGNAAHWSTKPPICQRIPCGLPPTIANGDFIS TNRENFHYGSVVTYRCNPGSGGRKVFELVGE
						PSIYCTSNDDQVGIWSGPAPQCIIPNKCTPPNV
						ENGILVSDNRSLFSLNEVVEFRCQPGFVMKGP
ļ					į	RRVKCQALNKWEPELPSCSRVCQPPPDVLHA
						ERTQRDKDNFSPGQEVFYSCEPGYDLRGAAS
}						MRCTPQGDWSPAAPTCEVKSCDDFMGQLLN
	}					GRVLFPVNLQLGAKVDFVCDEGFQLKGSSAS
						YCVLAGMESLWNSSVPVCEQIFCPSPPVIPNG
						RHTGKPLEVFPFGKAVNYTCDPHPDRGTSFD
ľ		ĺ			Ì	LIGESTIRCTSDPQGNGVWSSPAPRCGILGHC
		1				QAPDHFLFAKLKTQTNASDFPIGTSLKYECRP
						EYYGRPFSITCLDNLVWSSPKDVCKRKSCKTP
						PDPVNGMVHVITDIQVGSRINYSCTTGHRLIG
						HSSAECILSGNTAHWSTKPPICQRIPCGLPPTI
						ANGDFISTNRENFHYGSVVTYRCNLGSRGRK
i						VFELVGEPSIYCTSNDDQVGIWSGPAPQCIIPN
				·		KCTPPNVENGILVSDNRSLFSLNEVVEFRCQP
						GFVMKGPRRVKCQALNKWEPELPSCSRVCQ
						PPPEILHGEHTPSHQDNFSPGQEVFYSCEPGY
			}			DLRGAASLHCTPQGDWSPEAPRCAVKSCDDF
			f			LGQLPHGRVLFPLNLQLGAKVSFVCDEGFRL KGSSVSHCVLVGMRSLWNNSVPVCEHIFCPN
						PPAILNGRHTGTPSGDIPYGKEISYTCDPHPDR
					,	GMTFNLIGESTIRCTSDPHGNGVWSSPAPRCE
						LSVRAGHCKTPEQFPFASPTIPINDFEFPVGTS
						LNYECRPGYFGKMFSISCLENLVWSSVEDNC
	,		1			RRKSCGPPPEPFNGMVHINTDTQFGSTVNYSC
						NEGFRLIGSPSTTCLVSGNNVTWDKKAPICEII
[1			SCEPPPTISNGDFYSNNRTSFHNGTVVTYQCH
						TGPDGEQLFELVGERSIYCTSKDDQVGVWSS
						PPPRCISTNKCTAPEVENAIRVPGNRSFFSLTEI
						IRFRCQPGFVMVGSHTVQCQTNGRWGPKLPH
1					·	CSRVCQPPPEILHGEHTLSHQDNFSPGQEVFY
						SCEPSYDLRGAASLHCTPQGDWSPEAPRCTV
]			KSCDDFLGQLPHGRVLLPLNLQLGAKVSFVC
						DEGFRLKGRSASHCVLAGMKALWNSSVPVC
				1		EQIFCPNPPAILNGRHTGTPLGDIPYGKEVSYT
			· [ļ		CDPHPDRGMTFNLIGESTIRRTSEPHGNGVWS
				1		SPAPRCELPVGAACPHPPKIQNGHYIGGHVSL
	,	İ			ļ	YLPGMTISYTCDPGYLLVGKGFIFCTDQGIWS
1 1		j				QLDHYCKEVNCSFPLFMNGISKELEMKKVYH
	l			}		YGDYVTLKCEDGYTLEGSPWSQCQADDRWD
				ļ		PPLAKCTSRTHDALIVGTLSGTIFFILLIIFLSWI ILKHRKGNNAHENPKEVAIHLHSOGGSSVHP
]						RTLQTNEENSRVLP
803	2153	A	6574	2	3233	HGRSARLAAVPAEAMPGPRRPAGSRLRLLLL
""	2173	^	3574	-	ددعد	LLLPPLLLLLRG\SHAGNLTVAVVLPLANTSY
		-				PWSWA\RVGPAVELALAQVKARPDLLPGWT
		j	l			VRTVLGSSENALGVCSDTAAPLAAVDLKWE
		i				HNPAVFLGPGCVYAAAPVGRFTAHWRVPLL
			1			TAGAPALGFGVKDEYALTTRAGPSYAKLGDF
			ł			VAALHRRLGWERQALMLYAYRPGDEEHCFF
	.		İ		,	LVEGLFMRVRDRLNITVDHLEFAEDDLSHYT
}	}	- 1	ŀ			RLLRTMPRKGRVIYICSSPDAFRTLMLLALEA
	ļ]	ļ		GLCGEDYVFFHLDIFGQSLQGGQGPAPRRPW
	ŀ			1		ERGDGQDVSARQAFQAAKIITYKDPDNPEYL
						EFLKQLKHLAYEQFNFTMEDGLVNTIPASFH

- C	1 656					
SEQ ID NO: of	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	NO: of peptide	hod	ID NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine.
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	}			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan.
ļ			j	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
	ł	ļ	}	sequence		nucleotide insertion
						DGLLLYIQAVTETLAHGGTVTDGENITQRMW
						NRSFQGVTGYLKIDSSGDRETDFSLWDMDPE
			1			NGAFRVVLNYNGTSQELVAVSGRKLNWPLG
1			ļ	1		YPPPDIPKCGFDNEDPACNQDHLSTLEVLALV
1	1	!				GSLSLLGILIVSFFIYRKMQLEKELASELWRVR
	1	1				WEDVEPSSLERHLRSAGSRLTLSGRGSNYGSL
1		l	}	}	:	LTTEGQFQVFAKTAYYKGNLVAVKRVNRKR
	1					IELTRKVLFELKHMRDVQNEHLTRFVGACTD
ĺ		{			1	PPNICILTEYCPRGSLQDILENESITLDWMFRY
		l]	SLTNDIVKGMLFLHNGAICSHGNLKSSNCVV
		ļ				DGRFVLKITDYGLESFRDLDPEQGHTVYAKK
	İ					LWTAPELLRMASPPVRGSQAGDVYSFGIILQE
Ī		1				IALRSGVFHVEGLDLSPKEHERVTRGEQPPFR PSLALQSHLEELGLLMQRCWAEDPQERPPFQ
	1		1			QIRLTLRKFNRENSSNILDNLLSRMEQYANNL
						EELVEERTQAYLEEKRKAEALLYQILPHSVAE
ľ						QLKRGETVQAEAFDSVTIYFSDIVGFTALSAE
1	i		1		l	STPMQVVTLLNDLYTCFDAVIDNFDVYKVET
					, i	IGDAYMVVSGLPVRNGRLHACEVARMALAL
						LDAVRSFRIRHRPQEQLRLRIGIHTGPVCAGV
	1		1			VGLKMPRYCLFGDTVNTASRMESNGEAL\KI
			[HLSS\ETKAVL\EEFGGFELELRGDVEMKGKG
-	-					KVRTYWLLGERGSSTRG
804	2154	Α	6585	2	3837	DAPGRPPVRLPTMELEDGVVYQEEPGGSGAV
						MSERVSGLAGSIYREFERLIVRYDEEVVKELIP
						LVVAVLENLDSVFAQDQEHQVELELLRDDNE
						QLITQYEREKALRKHAEEKFIEFEDSQEQEKK DLQTRVESLESQTRQLELKAKNYADQISILEE
						REAELKKEYNALHQRHTEMIHNYMEHLERT
	}					KLHQLSGSDQLESTAHSRIRKERPISLGIFPLP
					·	AGDGLLTPDAQKGGETPGSEQWKFQELSQPR
						SHTSLKDELSDVSQGGSKATTPASTANSDVA
}						TIPTDTPLKEENEGFVKVTDAPNKSEISKHIEV
			.		·	QVAQETRNVSTGSAENEEKSEVQAHESTPEL
1	(DMDKDLSGYKGSSTPTKGIENKAFDRNTESL
			i			FEELSSAGSGLIGDVDEGADLLGMGREVENLI
ļ						LENTQLLETKNALNIVKNDLIAKVDELTCEK
1				ĺ	ĺ	DVLQGELEAVKQAKLKLEEKNRELEEELRKA
						RAEAEDARQKAKDDDDSDIPTAQRKRFTRVE
J]			ŀ		MARVLMERNQYKERLMELQEAVRWTEMIR
			+	ĺ	ſ	ASRENPAMQEKKRSSIWQFFSRLFSSSSNTTK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP
						GDKSKAFDFLSEETEASLASKREQKREQYRQ
				ļ	ļ	VKAHVQKEDGRVQAFGWSLPQKYKQVTNG
	}					QGENKMKNLPVPVYLRPLDEKDTSMKLWCA
						VGVNLSGGKTRDGGSVVGASVFYKDVAGLD
	1			ļ	j	TEGSKQRSASQSSLDKLDQELKEQQKELKNO
						EELSSLVWICTSTHSATKVLIIDAVOPGNILDS
		İ			ļ	FTVCNSHVLCIASVPGARETDYPAGEDLSESG
}	1				1	QVDKASLCGSMTSNSSAETDSLLGGITVVGC
			l			SAEGVTGAATSPSTNGASPVMDKPPEMEAEN
	[İ	SEVDENVPTAEE\ATEATEGNAGSAEDTV\DIS
1	1 1		}	ļ	1	QTGVYTEHVFTDPLG\VQIPEDLSPVYQSSND
1	i I		j			SDAYKDQISVLPNEQDLVREEAQKMSSLLPT
						MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD
1		1	l	- 1	1	SILSIVHVKGIVLVALADGTLAIFHRGVDGQW
I] 1	ļ		l	ļ	DLSNYHLLDLGRPHISIRCMTVVHDKVWCG
		ſ	[[ĺ	YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ
1						
l] [1	Į		LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ DVDIEPYVSKMLGTGKLGFSFVRITALMVSC

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion NRLWVGTGNGVIISIPLTETVILHQGRLLGLR
					\$ 5	ANKTSGVPGNRPGSVIRVYGDENSDKVTPGT FIPYCSMAHAQLCFHGHRDAVKFFVAVPGQV ISPQSSSSGTDLTGDKGRGHLHRSLVVRRP
805	2155	A	6605	469	2602	FGRLLWGTAFKSWKMKAPIPHLILLYATFTQ SLKVVTKRGSADGCTDWSIDIKKYQVLVGEP VRIKCALFYGYIRTNYSLAQSAGLSLMWYKS SGPGDFEEPIAFDGSRMSKEEDSIWFRPTLLQ DSGLYACVIRNSTYCMKVSISLTVGENDTGL CYNSKMKYFEKAELSKSKEISCRDIEDFLLPT REPEILWYKECRTKTWRPSIVFKRDTLLIREV REDDIGNYTCELKYGGFVVRRTTELTVTAPL TDKPPKLLYPMESKLTIQETQLGDSANLTCRA FFGYSGDVSPLIYWMKGEKFIEDLDENRVWE SDIKILKEHLGEQEVSISLIVDSVEEGDLGNYS CYVENGNGRRHASVLLHKRELMYTVELAGG LGAILLLLVCLVTIYKCYKIEIMLFYRNHFGA EELDGDNKDYDAYLSYTKVDPDQWNQETGE EERFALEILPDMLEKHYGYKLFIPDRDLIPTGT YIEDVARCVDQSKRLIIVMTPNYVVRRGWSIF ELETRLRNMLVTGEIKVILIECSELRGIMNYQE VEALKHTIKLLTVIKWHGPKCNKLNSKFWKR LQYEMPFKRIEPITHEQALDVSEQGPFGELQT VSAISMAAATSTALATAHPDLRSTFHNTYHS QMRQKHYYRSYEYDVPPTGTLPLTSIGNQHT YCNIPMTLINGQRPQTKSSREQNPDEAHTNSA ILPLLPRETSISSVIW
806	2156	A	6614	3	1584	NSARGGVGVRGARAMATVQEKAAALNLSAL HSPAHRPPGFSVAQKPFGATYVWSSINTLQT QVEVKKRRHRLKRHNDCFVGSEAVDVIFSHL IQNKYFGDVDIPRAKVVRVCQALMDYKVFE AVPTKVFGKDKKPTFEDSSCSLYRFTTIPNQD SQLGKENKLYSPARYADALFKSSDIRSASLED LWENLSLKPANSPHVNISTTLSPQVINEVWQE ETIGRLLQLVDLPLLDSLLKQQEAVPKIPQPK RQSTMVNSSNYLDRGILKAYSDSQEDEWLSA AIDCLEYLPDQMVVEISRSFPEQPDRTDLVKE LLFDAIGRYYSSREPLLNHLSDVHNGIAELLV NGKTEIALEATQLLLKLLDFQNREEFRRLLYF MAVAANPSEFKLQKESDNRMVVKRIFSKAIV DNKNLSKGKTDLLVLFLMDHQKDVFKIPGT L\HKIVS\VK\LMAIQNGRDPNRDAGYIYCQRI DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKK\LLGQFYKCHPDIFIEHFGD
807	2157	A	6615	4198	2094	FGIVGTFALETDELDSDRDPAIFSLCDFGAMR PQILLLALLTLGLAAQHQDKVPCKM/VKML CPDRVDKKVSCQVLGLLQVPSVLPPDTETLD LSGNQLRSILASPLGFYTALRHI.DLSTNEISFL QPGAFQALTHLEHLSLAHNRLAMATALSAG GLGPLPRVTSLDLSGNSLYSGLLERLLGEAPS LHTLSLAENSLTRLTRHTFRDMPALEQLDLHS NVLMDIEDGAFEGLPRLTHLNLSRNSLTCISD FSLQQLRVLDLSCNSIEAFQTAS\QPQAEFQLT WLDLRENKLLHFPDLAALPRLIYLNLSNNLIR LPTGPPQDSKGIHAPSEGWSALPLS\APSGNAS GRPLSQLLNLDLSYNEIELIPDSFLEHLTSLCFL NLSRNCLRTFEARRLGSLPCLMILDLSHNALE TLELGARALG\SLRTLLLQGNALRDLPPYTFA NLASLQRLNLQGNRVSPCGGPDEPGP\SGCV\ AFSGITSLRSLSLVDNEIELLRAGAFLHTPLTE

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \-possible nucleotide insertion LDLSSNPGLEVATGALGGLEASLEVLALQGN GLMVLQVDLPCFICLKRLNLAENRLSHLPAW TQAVSLEVLDLRNNSFSLLPGSAMGGLETSLR RLYLQGNPLSCCGNGWLAAQLHQGRVDVDA TQDLICRFSSQEEVSLSHVRPEDCEKGGLKNI NLIIILTFILVSAILLTTLAACCCVRRQKFNQQ
808	2158	A	6619	153	1852	FKALSQYIYTNTHLEREAAFEVAILLRRMEEG ARHRNNTEKKHPGGGESDASPEAGSGGGV ALKKEIGLVSACGIIVGNIIGSGIFVSPKGVLEN AGSVGLALIVWIVTGFITVVGALCYAELGVNI PKSGGDYFYVKDIFGGLAGFLRLWIAVLVIYP TNQAVIALTFSNYVLQPLFPTCFPPESGLRLLA AICLLLLTWVNCSSVRWATRVQDIFTAGKLL ALALIIIMGIVQICKGEYFWLEPKNAFENFQEP DIGLVALAFLQGSFAYGGWNFLNYVTEELV DPYKNLVPRAIFISIPLVTFVYVFANV/ALYVT AMSPQEL/LAS/NAVAVTFGEKLLGVMAWIM PISVALSTFGGVNGSLFTSSRLFFAGAREGHLP SVLAMIHVKRCTPIPALLFTCISTLLMLVTSD MYTLINYVGFINYLFFYGVTVAGQIVLRWKKP DIPRPIKINLLFPIIYLLFWAFLLVFSLWSEPVV CGIGLAIMLTGVPVYFLGVYWQHKPKCFSDFI ELLTLVSQKMCVVVYPEVERGSGTEEANED MEEQQQPMYQPTPTKDKDVAGQPQP
809	2159	A	6621	1041	223	QDSRKMLPSTSVNSLVQGNGVLNSRDAARH TAGAKRYKYLRRLFRFRQMDFEFAAWQMLY LFTSPQRVYRNFHYRKQTKDQWARDDPAFL VLLSIWLCVSTIGFGFVLDMGFFETIKLLLWV VLIDCVGVGLLIATLMWFISNKYLVKRQSRD YDVEWGYAFDVHLNAFYPLLVILHFIQLFFIN
						HVILTDTFIGYLVONTLWLVAVGYYYYVTFL GYSVGLLFFS\ALPFLKNTVILLYPFAPLILLYG
810	2160	A	6623	160	822	LSLALGWNFTHTLCSFYKYRVK SPASGHCRLNGAAVAMFGCLVAGRLVQTAA QQVAEDKFVFDLPDYESINHVVVFMLGTIPFP EGMGGSVYFSYPDSNGMPVWQLLGFVTNGK PSAIFKISGLKSGEGSQHPFGAMNIVRTPSVAQ IGISVELLDSMAQQTPVGNAAVSSVDSFTQFT QKMLDNFYNFASSFAVSQ/VPDDTQ/RPSEMF IPANVVLKWYENFQRRTSTEPSLLENIIWIKIN F
811	2161	A	6627	18	3367	LEGSLNTERAKYYLTITMPHFTVTKVEDPEEG AAASISQEPSLADIKARIQDSDEPDLSQNSITG EHSQLLDDGHKKARNAYLNNSNYEEGDEYF DKNLALFEEEMDTRPKVSSLLNRMANYTNLT QGAKEHEEAENITEGKKKPTKTPQMGTFMG VYLPCLQNIFGVILFLRLTWVVGTAGVLQAF AIVLICCCCTMLTAISMSAIATNGVVPAGGSY FMISRALGPEFGGAVGLCFYLGTTFAAAMYIL GAIEIFLVYIVPRAAIFHSDDALKESAAMLNN MRVYGTAFLVLMVLVVFIGVRYVNKFASLFL ACVIVSILAIYAGAIKSSFAPPHFPVCMLGNRT LSSRHIDVCSKTKEINNMTVPSKLWGFFCNSS QFFNATCDEYFVHNNVTSIQGIPGLASGIITEN LWSNYLPKGEIIEKPSAKSSDVLGSLNHEYVL VDITTSFTLLVGIFFPSVTGIMAGSNRSGDLKD AQKSIPIGTILAILTTSFVYLSNVVLFGACIEGV VLRDKFGDAVKGNLVVGTLSWPSPWVIVIGS FFSTCGAGLQSLTGAPRLLQAIAKDNIIPFLRV

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first am ino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion FGHSKANGEPTWALLTAAIAELGILIASLDL VAPILSMFFLMCYLFVNLACALQTLLRTPNW RPRFRYYHWALSFMGMSICLALMFISSWYYA IVAMVIAGMIYKYIEYQGAEKEWGDGIRGLS LSAARFALLRLEEGPPHTKNWRPQLLVLLKL DEDLHVKHPRLLTFASQLKAGKGLTVGSVIV GNFLENYGEALAAEQTIKHLMEAEKVKGFCQ LVVAAKLREGISHLIQSCGLGGMKHNTVVM GWPNGWRQSEDARAWKTFIGTVRVTTAAHL ALLVAKNISFFPSNVEQFSEGNIDVWWIVHDG GMLMLLPFLLK\QHKVWRKCSIRFF\TVAQLE DNSIQMKKDLATFLYHLRIEAEVEVVEMHDS DISAYTYERTLMMEQRSQMLRHMRLSKTER DREAQLVKDRNSMLRLTSIGSDEDEETETYQ EKVHMTWTKDKYMASRGQKAKSMEGFQDL LNMRPDQSNVRRMHTAVKLNEVIVNKSHEA KLVLLNMPGPPRNPEGDENYMEFLEVLTEGL ERVLLVRGGGSEVITIYS
812	2162	A	6628	66	640	AVCTMSEMAELSELYEESSDLQMDVMPGEG DLPQMEVGSGSRELSLRPSRSGAQQLEEEGP MEEEEAQPMAAPEGKRSLANGPNAGEQPGQ VAGADFESEDEGEEFDDWEDDYDYPEEEQLS GAGYRVSAALEEADKMFLRTREPALDGGFQ MHYEKTPFDQLAFIEELF\SLMVVNRLTEELG CDEIIDRE
813	2163	A	6630	708	1355	AKMGAYKYIQELWRKKQSDVMRFLLRVRC WQYRQLSALHRAPRPTRPDKARRLGYKAKQ GY/VYIYIGFVFAVIYRIRVRRGGRKRPVPKG ATYGKPVHHGVNQLKFARSLQSVAEERAGR HCGALRVLNSYWVGEDSTYKFFEVILIDPFHK AIRRNPDTQWITKPVHKHREMRGLTSAGRKS RGLGKGHKFHHTIGGSRRAAWRRNTLQLH RYR
814	2164	A	6635	201	1705	KGTEMNKSR WQSRRRHGRRSHQQNPWFRLR DSEDRSDSRAAQPAHDSGHGDDESPSTSSGT AGTSSVPELPGFYFDPEKKRYFRLLPGHNNCN PLTKESIRQKEMESKRLRLLQEEDRRKKIARM GFNASSMLRKSQLGFLNVTNYCHLAHELRLS CMERKKVQIRSMDPSALASDRFNLILADTNS DRLFTVNDVTVGGSKYGIINLQSLKTPTLKVF MHENLYFTNRKV\NSVCWASLNHLDSHILLC LMGLAETPGCATLLPASLFVNSHPAGIDRPG\ MLCSFRIPGAWSCAWSLNIQANNCFSTGLSR RVLLTNVVTGHRQSFGTNSDVLAQQFALMA PLLFNGCRSGEIFAIDLRCGNQGKGWKATRLF HDSAVTSVRILQDEQYLMASDMAGKIKLWD LRTTKCVRQYEGHVNEYAYLPLHVHEEEGIL VAVGQDCYTRIWSLHDARLLRTIPSPYPASKA DIPSVAFSSRLGGSRGAPGLLMAVGQDLYCY SYS
815	2165	A	6643	659	3282	NKNILEVPSARTTRIMGDHLDLLLGVVLMAG PVFGIPSCSFDGRIAFYRFCNLTQVPQVLNTTE RLLLSFNYIRTVTASSFPFLEQLQLLELGSQYT PLTIDKEAFRNLPNLRILDLGSSKIYFLHPDAF QGLFHLFELRLYFCGLSDAVLKDGYFRNLKA LTRLDLSKNQIRSLYLHPSFGKLNSLKSIDFSS NQIFLVCEHELEPLQGKTLSFFSLAANSLYSR VSVDWGKCMNPFRNMVLEILDVSGNGWTV DITGNFSNAISKSQAFSLILAHHIMGAGFGFHN IKDPDQNTFAGLARSSVRHLDLSHGFVFSLNS

[CFA ***	0000	T .:-	. 000	I Due !	I S. W. C.	
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of peptide	hod	ID NO:	beginning nucleotide	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
1	1		in		location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		ſ		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
			<u>}</u>	sequence		nucleotide insertion
! !			1	<u> </u>	}	RVFETLKDLKVLNLAYNKINKIADEAFYGLD
		l			i	NLQVLNLSYNLLGELYSSNFYGLPKVAYIDL
1		1				QKNHIAIIQDQTFKFLEKLQTLDLRDNALTTIH
		1				FIPSIPDIFLSGNKLVTLPKINLTANLIHLSENR
		1				LENLDILYFLLRVPHLQILILNQNRFSSCSGDQ
1 1	[1	[ĺ	TPSENPSLEQLFLGENMLQLAWETELCWDVF
					[EGLSHLQVLYLNHNYLNSLPPGVFSHLTALR
1 !	<u> </u>					GLSLNSNRLTVLSHNDLPANLEILDISRNQLL
						APNPDVFVSLSVLDITHNKFICECELSTFINWL
1 1	1	1	i	1	}	NHTNVTIAGPPADIYCVYPDSLSGVSLFSLSTE
1 1]			ĺ		GCDEEEVLKSLKFSLFIVCTVTLTLFLMTILTV
1 1						TKFRGFCFICYKTAQRLVFKDHPQGTEPDMY
						KYDAYLCFSSKDFTWVQNALLKHLDTQYSD
j:		J]	[QNRFNLCFEERDFVPGENRP\ANIQDAIWNSR
		ļ		(:		KIVCLVSRHFLRDGWCLEAFSYAQGRCLSDL
						NSALIMVVVGSLSQYQLMKHQSIRGFVQKQQ
						YLRWPEDLQDVGWFLHKLSQQILKKEKEKK
						KDNNIPLQTVATIS
816	2166	Α	6646	1	3811	RDRAGVRPAGKQHAAAAFYDVGGDRPWDS
				ļ		GNTQLPPRNPVKANAMFGAGDEDDTDFLSPS
				Ì		GGARLASLFGLDQAAAGHGNEFFQYTAPKQP
1	İ			İ		KKGQGTAATGNQATPKTAPATMSTPTILVAT
1 1				ļ	}	AVHAYRYTNGQYVKQGKFGAAVLGNHTTR
j		i				EYRILLYISQQQPVTVARIHVNFELMVRPNNY
1						STFYDDQRQNWSIMFESEKAAVEFNKQVCIA
1					'	KCNSTSSLDAVLSQDLIVADGPAVEVGDSLE
					,	VAYTGWLFQNHVLGQVFDSTANKDKLLRLK
		[LGSGKVIKGWEDGMLGMKKGGKRLLIVPPA
		ŀ				CAVGSEGVIGWTQATDSILVFEVEVRRVKIA
						KDSGSDGHSVSSRDSAAPSPIPGADNLSADPV
\vdash						VSPPTSIPFKSGEPALRTKSNSLSEQLAINTSPD
1 1		ł	ł			AVKAKLISRMAKMGQPMLPILPPQLDSNDSEI
						EDVNTLQGGGQPVVTPSVQPSLQPAHPALPQ
1						MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA
						VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ
J I		ŀ	İ			PVRPLYPAPLSQPPHFQGSGDMASFLMTEAR
						QHNTEIRMAVSKVADKMDHLMTKVEELQKH
			1		· ·	SAGNSMLIPSMSVTMETSMIMSNIQRIIQENER
						LKQEILEKSNRIEEQNDKISELIERNQRYVEQS
						NLMMEKRNNSLQTATENTQARVLHAEQEKA
		1	1	ľ]	KVTEELAAATAQVSHLQLKMTAHQKKETEL
				i		QMQLTESLKETDLLRGQLTKVQAKLSELQET
1 1						SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL
l . l						RVEKESLEKNLSERKKKSAQERSQAEEEIDEI
	l .					RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS
				•		LVQAELQTQWEAKCEHLLASAKDEHLQQYQ
						EVCAQRDAYQQKLVQLQEKSVCFA\CLALQA
						QITALTKQNEQHIKELEKNKSQMSGVEAAAS
						DPSEKVKKIMNQVFQSLRREFELEESYNGRTI
		1			ľ	LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE
					i	EKAEERPRRPSQEQSASASSGQPQAPLNRERP
		1				ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR
						KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP
, !					ł	TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN
1 1						PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE
[[I	i	
						GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK
				,		GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK
						GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide delction, \=possible nucleotide insertion KS/ARNSQLRIVLVGKTGAGKSATGNSILGRK VFHSGTAAKSITKKCEKRSSSWKETELVVVD TPGIFDTEVPNAETSKEIIRCILLTSPGPHALLL VVPLGRYTEEEHKATEKILKMFGERARSFMIL IFTRKDDLGDTNLHDYLREAPEDIQDLMDIFG DRYCALNNKATGAEQEAQRAQLLGLIQRVV RENKEGCYTNRMYQRAEEEIQKQTQAMQEL HRVELEREKARIREEYEEKIRKLEDKVEQEKR KKQMEKKLAEQEAHYAVRQQRARTEVESKD GILELIMTALQIASFILLRLFAED
818	2168	A	6660	357	1890	APSGSWTRVVLTLDPCSLRSRSPRSLLDPGMP GISARGLSHEGRKQLAVNLTRVLALYRSILDA YIIEFF\TDNLWDTLPCSWQEALDGLKPPQLA TMLLGMPGEGEVVRYRSVWPLTLLALKSTA CALAFTRMPGFQTPSEFLENPSQSSRLTAPFR KHVRPKKQHEIRRLGELVKKLSDFT/GLHPGC RRGLRPG\HLSRFMALGLGLMVKSIEGDQRL VERAQRLDQELLQALEKEEKRNPQVVQTSPR HSPHHVVRWVDPTALCEELLLPLENPCQGRA RLLLTGLHACG\DLSVALLRHFSCCPEVVALA SVGCCYMKLSDPGGYPLSQWVAGLPGYELP YRLREGACHALEEYAERLQKAGPGLRTHCY RAALETVIRRARPELRRPGVQGIPRVHELKIEE YVQRGLQRVGLDPQLPLNLAALQAHLAQEN RVVAFFSLALLLAPLVETLILLDRLLYLQEQA LSP\GFHAELLPIFSPELSPRNLVLVATKMPLG QALSVLETEDS
819	2169	A	6661	. ·	2686	SGSGHCLAEAASMGPWGWKLRWTVALLLA AAGTAVGDRCERNEFQCQDGKCISYKWVCD GSAECQDGSDESQETCLSVTCKSGDFSCGGR VNRCIPQFWRCDGQVDCDNGSDEQGCPPKTC SQDEFRCHDGKCISRQFVCDSDRDCLDGSDE ASCPVLTCGPASFQCNSSTCIPQLWACDNDPD CEDGSDEWPQRCRGLYVFQGDSSPCSAFEFH CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH TKACKAVGSIAYLFFINRHEVRKMTLDRSEY TSLIPNLRNVVALDTEVASNRIYWSDLSQRMI CSTQLDRAHGVSSYDTVISRDIQAPDGLAVD WIHSNIYWTDSVLGTVSVADTKGVKRKTLFR ENGSKPRAIVVDPVHGFMYWTDWGTPAKIK KGGLNGVDIYSLVTENIQWPNGITLDLLSGRL YWVDSKLHSISSIDVNGGNRKTILEDEKRLAH PFSLAVFEDKVFWTDIINEAIFSANRLTGSDV NLLAENLLSPEDMVLFHNLTQPRGVNWCERT TLSNGGCQYLCLPAPQINPHSPKFTCACPDGM LLAR/DMRSCLTEG/EAAVATQETSTVRLKVS STAVRTQHTTTRPVPDTSRLPGATPGLTTVEI VTMSHQALGDVAG/RGN/EKKPSSVRALSIVL PIVLLVFLCLGVFLLWKNWRLKNINSINFDNP VYQKTTEDEVHICHNQDGYSYPSRQMVSLED DVA
820	2170	A	6666	17	4146	ERGISSQIKGMKSGSGGGSPTSLWGLLFLSAA LSLWPTSGEICGPGIDIRNDYQQLKRLENCTVI EGYLHILLISKAEDYRSYRFPKLTVITEYLLLF RVAGLESLGDLFPNLTVIRGWKLFYNYALVIF

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
					:	EMTNLKDIGLYNLRNITRGAIRIEKNADLCYL STVDWSLILDAVSNNYIVGNKPPKECGDLCP GTMEEKPMCEKTTINNEYNYRCWTTNRCQK MCPSTCGKRACTENNECCHPECLGSCSAPDN DTACVACRHYYYAGVCVPACPPNTYRFEGW RCVDRDFCANILSAESSDSEGFVIHDGECMQE CPSGFIRNGSQSMYCIPCEGPCPKVCEEKKT KTIDSVTSAQMLQGCTIFKGNLLINIRRGNNIA SELENFMGLIEVVTGYVKIRHSHALVSLSFLK NLRLILGEEQLEGNYSFYVLDNQNLQQLWD WDHRNLTIKAGKMYFAFNPKLCVSEIYRMEE VTGTKGRQSKGDINTRNNGERASCESDVLHF TSTTTSKNRIIITWHRYRPPDYRDLISFTVYYK EAPFKNVTEYDGQDACGSNSWNMVDVDLPP NKDVEPGILLHGLKPWTQYAVYKAVTLTM VENDHIRGAKSEILYIRTNASVPSIPLDVLSAS NSSSQLIVKWNPPSLPNGNLSYYIVRWQRQP QDGYLYRHNYCSKDKIPIRKYADGTIDIEEVT ENPKTEVCGGEKGPCCACPKTEAEKQAEKEE AEYRKVFENFLHNSIFVPRPEKRRDVMQVA
					: :	NITMSSRSRNITAADTYNITDPEELETEYPFF ESRVDNKERTVISNLRPFTLYRIDIHSCNHEAE KLGCSASNFVFARTMPAEGADDIPGPVTWEP RPENSIFLKWPEPENPNGLILMYEIKYGSQVE DQRECVSRQEYRKYGGAKLNRLNPGNYTARI QATSLSGNGSWTDPVFFYVQAKRYENFIHLII ALPVAVLLIVGGLVIMLYVFHKKRNNSRLGN GVLYASVNPEYFSAADVYVPDEWEVAREKIT MSRELGQGSFGMVYEGVAKGVVKDEPETRV AIKTVNEAASMRERIEFLNEASVMKEFNCHH VVRLLGVVSQGQPTLVIMELMTRGDLKSYLR
821	2171				,	SLRPEMENNPVLAPPSLSKMIQMAGEIADGM AYLNANKFVHRDLAARNCMVAEDFTVKIGD FGMTRDIYETDYYRKGGKGLLPVRWMSPESL KDGVFTTYSDVWSFGVVLWEIATLAEQPYQ GLSNEQVLRFV\MEGGLLDKPDNCPDMLFEL MRMCWQYNPKMRPSFLEIISSIKEEMEPGFRE VSFYYSEENKLPEPEELDLEPENMESVPLDPS ASSSSLPLPDRHSGHKAENGPGPGVLVLRASF DERQPYAHMNGGRKNERALPLPQSSTC
021	2171	A	6691	106	825	GRVLFRGCGVGHKGQVLMGTFILAQDWLSE SNHVFCVSSMLRLQKRLASSVLRCGKKKVW LDPNETNEIANANSRQQIRKLIKDGLIIRKPVT VHSRARCRKNTLARRKGRHMGIGKRKGTAN ARMPEKVTWMRRMRILRRLLRYRES/KRYR ESKKIDRHMYHSLYLKVKGNVFKNKRILMEH IHKLKADKARKKLLADQAEARRSKTKEARK RREERLQAKKEEIIKTLSKEEETKK
822	2172	A .	6715	772	,	DFRPGLLLPRKKKMFGFHKPKMYRSIEGC\CI SGAKSSS\RFTDSKRYEK\DFQ\SCFGLHETR\ SGDI\CNA\CVLL\LKRWKKLPAGSKK\NWNH VVDARAGPS\LKTTLKPKKVKTL\SGNRIK\ST QISKLQKEFKR\HNSDAHSTTS\SASP\AQSPLF TVNQFRWTGSDTGVGFPGSNRNHPVFSFLDL\ TYWKRQKICCG\IYKGRFGEVLIDTHLFKPCC SNKKA\AAEKPEEQGPEPLPISTQEWVTEVFM
823	2173	Ą	6727	3	4063	PYLATLQLDSSLLIPPKYQTPPAAAQGQATPG NAGPLAPNGSAAPPAGSAFNPTSNSSSTNPAA SSSASGSSVPPVSSSASAPGISQISTTSSSGFSGS VGGQNPSTGGISADRTQGNIGCGGDTDPGQS

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl- eotide	peptide seq-		USSN	nucleotide location	location corresponding	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			•	peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
				,		SSQPSQDGQESNVPSVGSLADPDYLNTPQMN TRVTLNSA A BA SNIGG A GNI BORA TRDDEGNIERR
						TPVTLNSAAPASNSGAGVLPSPATPRFSVPTP RTPRTPRTPRGGGTASGOGSVKYDSTDOGSP
	 					ASTPSTTRPLNSVEPATMQPIPEAHSLYVTLIL
						SDSVMNIFKDRNFDSCCICACNMNIKGADVG
						LYIPDSSNEDQYRCTCGFSAIMNRKLGYNSGL
	<u> </u>					FLEDELDIFGKNSDIGQAAERRLMMCQSTFL
						PQVEGTKKPQEPPISLLLLLQNQHTQPFASLN
						FLDYISSNNRQTLPCVSWSYDRVQADNNDY
	ĺ					WTECFNALEQGRQYVDNPTGGKVDEALVRS
						ATVHSWPHSNVLDISMLSSQDVVRMLLSLQP FLQDAIQKKRTGRTWENIQHVQGPLTWQQFH
						KMAGRGTYGSEESPEPLPIPTLLVGYDKDFLT
						ISPFSLPFWERLLLDPYGGHRDVAYIVVCPEN
			1			EALLEGAKTFFRDLSAVYEMCRLGOHKPICK
1					İ	VLRDGIMRVGKTVAQKLTDELVSEWFNQPW
						SGEENDNHSRLKLYAQVCRHHLAPYLATLQL
						DSSLLIPPKYQTPPAAAQGQATPGNAGPLAPN
						GSAAPPAGSAFNPTSNSSSTNPAASSSASGSSV
						PPVSSSASAPGISQISTTSSSGFSGSVGGQNPST
						GGISADRTQGNIGCGGDTDPGQSSSQPSQDG QESVTERERIGIPTEPDSADSHAHPPAVVIYM
						VDPFTYAAEEDSTSGNFWLLSLMRCYTEMLD
						NLPEHMRNSFILQIVPCQYMLQTMKDEQVFY
						IQYLKSMAFSVYCQCRRPLPTQIHIKSLTGFGP
f						AASIEMTLKNPERPSPIQLYSPPFILAPIKDKQT
						ELGETFGEASQKYNVLFVGYCLSHDQRWLL
						ASCTDLHGELLETCVVNIALPNRSRRSKVSAR
1					-	KIGLQKLWEWCIGIVQMTSLPWRVVIGRLGR
1						LGHGELKDWSILLGECSLQTISKKLKDVCRM CGISAADSPSILSACLVAMEPQGSFVVMPDAV
						TMGSVFGRSTALNMQSSQLNTPQDASCTHIL
		J				VFPTSSTIQVAPANYPNEDGFSPNNDDMFVDL
						PFPDDMDNDIGILMTGNLHSSPNSSPVPSPGSP
						SGIGVGSHFQHSRSQGERLLSREAPEELKQQP
						LALGYFVSTAKAENLPQWFWSSCPQAQMQC
						PLFLKASLHHHISVAQTDELLPARNSQRVPHP
						LDSKTTSDVLRFVLEQYNALSWLTCNPATQD RTSCLPVHFVVLTQLYNAIMNIL
824	2174	A	6732	2440	365	VEEGLGRRRTPPGGRRGPVTPARPGPDSVRR
"	-11T	^	3,32	_,,,	303	RLLPPSSAAAFSSHRHNLLCSRRRGGGGGG
						GGGGGTIKRPGITGPTAATSPSGEPGNAASAP
		l	ľ	ĺ	İ	LSLLSPFPGQTTYQHPGVAEPSAYGGRDVAC
		l				ASLVFGRLQHRGGDRKRGLLGRSSGDAASD
		ľ				QPFRCRSGSTAGRLVKQMDFTEAYADTCSTV
		ĺ	Í	ĺ	ĺ	GLAAREGNVKVLRKLLKKGRSVDVADNRG
						WMPIHEAAYHNSVECLQMLINADSSENYIKM KTEEGECALHI AASOGHWKIVOULEAGADB
		1		İ		KTFEGFCALHLAASQGHWKIVQILLEAGADP NATTLEETTPLFLAVENGQIDVLRLLLQHGAN
						VNGSHSMCGWNSLHQASFQENAEIIKLLLRK
		1		ľ	ĺ	GANKECQDDFGITPLFVAAQYG\KLESL\SILIS
						SG\ANVNCQALDKATPLFIAAQEGHTKCVELL
						LSSGADPDLYCNEDSWQLPIHAAAQMGHTKI
					ļ	LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE
	1	i				DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ
		İ				KDCEFFGIVNILLKYGAQINELHLAYCLKYEK
		ļ				FSIFRYFLRKGCSLGPWNHIYEFVNHAIKAQA KYKEWLPHLLVAGFDPLILLCNSWIDSVSIDT
		l				LIFTLEFTNWKTLAPAVERMLSARASNAWIL
}	ļ		}		ļ	QQHIATVPSLTHLCRLEIRSSLKSERLRSDSYIS
L				1		

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion QLPLPRSLHNYLLYEDVLRMYEVPELAAIQD
825	2175	A	6735	277	1252	G RIMGLFDRGVQMLLTTVGAFAAFSLMTIAVG TDYWLYSRGVCKTKSVSENETSKKNEEVMT HSGLWRTCCLEGNFKGLCKQIDHFPEDADYE ADTAEYFLRAVRASSIFPILSVILLFMGGLCIA ASEFYKTRHNIILSAGIFFVSAGLSNIIGIIVYIS ANAGDPSKSDSKKNSYSYGWSFYFGALSFIIA EMVGVLAVHMFIDRHKQLRATARA\TDYLQ ASAITRIPSYRYRYQRRSRSSRSTEPSHSRDA SPVGIKGFNTLPSTEISMYTLSRDPLKAATTPT ATYNSDRDNSFLQVHNCIQKENKDSLHSNTA NRRTTPV
826	2176	A	6744	3	5177	SDDLRTGLFQDVQDAESLKLPGVYEVLFYNE TEDCPGMMLWRYPEPRGLTLVRITPVPFNTT EDPDISTADLGDVLQDPCSLEYWDELQKVFV AFREFNLSESKVCELQLPDINLVNDQKKLVSS DLWRIVLNSSQNGADDQSSASESGSQSTCDPL VTPTALAACTRVDSCFTPWFVPSLCVSFQFAH LEFHLCHHLDQLGTAAPQYLQPFVSDRNMPS ELEYMIVSFREPHMYLRQWNNGSVCQEIQFL AQADCKLLECRNVTMQSVVKPFSIFGQMAVS SDVVEKLLDCTVIVDSVFVNLGQHVVHSLNT AIQAWQQNKCPEVEELVFSHFVICNDTQETL RFGQVDTDENILLASLHSHQYSWRSHKSPQL LHICIEGWGNWRWSEPFSVDHAGTFIRTIQYR GRTASLIIKVQQLNGVQKQIIICGRQIICSYLSQ SIELKVVQHYIGQDGQAVVREHFDCLTAKQK LPSYILENNELTELCVKAKGDEDWSRDVCLE SKAPEYSIVIQVPSSNSSIIYVWCTVLTLEPNS QVQQRMIVFSPLFIMRSHLPDPIIIHLEKRSLGL
						SETQIIPGKGQEKPLQNIEPDLVHHLTFQAREE YDPSDCAVPISTSLIKQIATKVHPGGTVNQILD EFYGPEKSLQPIWPYNKKDSDRNEQLSQWDS PMRVKLSIWKPYVRTLLIELLPWALLINESKW DLWLFEGEKIVLQVPAGKIIIPPNFQEAFQIGIY WANTNTVHKSVAIKLVHNLTSPKWKDGGNG EVVTLDEEAFVDTEIRLGAFPGHQKLCQFCIS SMVQQGIQIIQIEDKTTIINNTPYQIFYKPQLSV CNPHSGKEYFRVPDSATFSICPGGEQPAMKSS SLPCWDLMPDISQSVLDASLLQKQIMLGFSPA PGADSSQCWSLPAIVRPEFPRQSVAVPLGNFR ENGFCTRAIVLTYQEHLGVTYLTLSEDPSPRV IIHNRCPVKMLIKENIKDIPKFEVYCKKIPSECS IHHELYHQISSYPDCKTKDLLPSLLLRVEPLDE VTTEWSDAIDINSQGTQVVFLTGFGYVYVDV VHQCGTVFITVAPEGKAGPILTNTNRAPEKIV TF/KMPITQLSLAVFDDLTHHKASAELLRLTL DNIFLCVAPGAGPLPGEEPVAALFELYCVEIC CGDLQLDNQLYNKSNFHFAVLVCQGEKAEPI QCSKMQSLLISNKELEEYKEKCFIKLCITLNEG KSILCDINEFSFELKPARLYVEDTFVYYIKTLF DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPVKLRKLVIQPVNLLVSIHASLKLYI ASDHTPLSFSVFERGPIFTTARQLVHALAMHY AAGALFRAGWVVGSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSGVSRGTTSFVK HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQLPESLGEGLRQGLSRLGISLLGAI

SEQ ID SEQ ID Met SEQ Predicted Predicted end Amino acid sequence (A=Alanine C=Cys NO: of NO: of NO: of nucl- peptide ID NO: beginning nucleotide location F=Phenylalanine, G=Glycine, H=Histidin	•
nucl- peptide in nucleotide location F=Phenylalanine, G=Glycine, H=Histidin	
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	e,
eotide seq- USSN location corresponding I=Isoleucine, K=Lysine, L=Leucine,	
seq- uence 09/496 correspondi to last amino M=Methionine, N=Asparagine, P=Proline	,
uence 914 ng to first acid residue Q=Glutamine, R=Arginine, S=Serine,	
amino acid of peptide T=Threonine, V=Valine, W=Tryptophan,	
residue of sequence Y=Tyrosine, X=Unknown, *=Stop codon	,
peptide /=possible nucleotide deletion, \=possible	
sequence nucleotide insertion	
VGKGIMGVFTKPIGGAAELVSQTGY	
GLSQLPKQRHQPSD\VHADQAPNSHY	
KMLQSLGRPEVHMALDVVLVRGSG	
LLLTSEVLFVVSVSEDTQQQAFPVTE	
SKQNNLLTVQLKQPRVACDVEVDGV	
QQYNRLVDYITKTSCHLAPSCSSMQI	
AEPPPSTVKTYHYLVDPHFAQVFLSK	FIMVK
NKALRKGFP 827 2177 A 6748 2 1662 FVGAPRRGNPFGSPGNPGRHQGPCHL	DD CTT
ASGVSPTLWRPQAAATGLEMPSSGR	
LDSGSLTSLDSSVFCSEGEGEPLALGI	
VGGSRFVLSQQALSCFPHTRLGKLAV RRPGALAAVPSPLELCDDANPVDNE	
SQAFRYVLHYYRTGRLHVMEQLCAI	
	•
QYWGIDELSIDSCCRDRYFRRKELSE DTEDQESQHESEQDFSQGPCPTVRQK	
EKPGSSTAARIFGVISIIFVGVSIINMA	
SWLDLQLLEILEYVCISWFTGEFVLRI	
RCRFLRKVPNIIDLLAILPFYITLLVES	
TQEL\ENVGAHCPGCLRLLRAL\RML	-
HSTGLRSLGMTITOCYEEVGLLLLFL	
STVEYFAEQSIPDTTFTSVPCAWWW	
TVGYGDIRPDTTTGKIVAFMCILSGIL	
AIINDRFSACYFTLKLKEAAVRQREA	
NIATDSYISVNLRDVYARSIMEMLRL	
ASTRSSGGDDFWF	·······
828 2178 A 6786 5672 1360 GTHPASSGPVPLPPAAVSAATREELG	EPVPFV
TASSGFQSMHSSNPKVRSSPSGNTQS	
EVMVRPPTVMSPSGNPQLDSKFSNQ	
ASQSQPSPCDSKSGGHTPKALPGPGG	
NGAGNGAKGKGKRERSISADSFDQR	
DDSDIKECNSADHIKSQDSQHTPHSM	
APRSSTPPHGQTTATEPTPAQKTPAK	/VYVFS
TEMANKAAEAVLKGQVETIVSFHIQN	ISNNK
TERSTAPLNTQISALRNDPKPLPQQPP	APANQ
DQNSSQNTRLQPTPPIPAPAPKPAAPP	
SPGVENKLIPSVGSPASSTPLPPDGTG	NSTPN
NRAVTPVSQGSNSSSADPKAPPPPPVS	
LGENPDGLSQEQLEHRERSLQTLRDIG)R MLFP
DEKEFTGAQSGGPQQNPGVLDGPQK	
QAMMAQSQSLGKGPGPRTDVGAPFO	
DVPFSPDEMVPPSMNSQSGTIGPDHLI	
EQIAWLKLQQEFYEEKRRKPEQVVV	
DMMVHQHGPRGVVRGPPPPYQMTPS	
, GGTEPFSDGINMPHSLPPRGMAPHPN	
MRLPGFAGMINSEMEGPNVPNPASRI	
SWPDDVPKIPDGRNFPPGQGIFSGPGF	
NPQGLSEEMFQQQLAEKQLGLPPGM.	
PSMEMNRMIPGSQRHMEPGNNPIFPR	
LSPSRGDFPKGIPPQMGPGRELEFGM	/PSGM
KGDVNLNVNMGSNSQMIPQKMREAG	
MLKLRPGGSDMLPAQQKMVPLPFGE	
YGMGPRPFLPMSQGPGSNSGLRNLRI	- 1
RTNSRLSHMPPLPLNPSSNPTSLNTAP	
LGRKPLDISVAGSQVHSPGINPLKSPT	
SPMLGSPSGNLKSPQTPSQLAGMLAG	
SIKSPPVLGSAAASPVHLKSPSLPAPSI	
PEPPLQSPGIPPNHKAPLTMASPAMLQ	1
GPPPPTASQPASVNIPG\SLPSSTPYTM SQNPLSIM\MSR\MSKFAM\PS\SNPGY	
	NHDAL L

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion KTVASSDDDSPPARSPNLPSMNNMPGMGINT QNPRISGPNPVVPMPTLSPMGMTQPLSHSNQ MPSPNAVGPNIPPHGVPMGPGLMSHNPIMGH GSQEPPMVPQGRMGFPQGFPPVQSPPQQVPFP HNGPSGQQGSFPGGMGFPGEGPLGRPSNLPQ SSADAALCKPGGPGGPDSFTVLGNSMPSVFT DPDLQEVIRPGATGIPEFDLSRIIPSEKPSQTLQ
829	2179	A	6797	433	3	YFPRGEVPGRKQPQGPGPGFSHMQGMMGEQ APRMGLALPGMGGPGPVGTPDIPLGTAPSMP GHNPMRPPAFLQQGMMGPHHRMMSPAQST MPGQPTLMSNPAAAVGMIPGKDRGPAGLYT HPGPVGSPGMMMSMQGMMGPNRTS ASFFNFSICICKIILEVGPPVGHPAHDDVGGRH GPGGR/GSRSPRSLQCAPGGGRRSGCPAGSSP
830	2180	A	i 6800	3	1911	ASTCPPSPGGSGADRFGPSPPPPSREAAPTAG AAASSTSSGASCPPVPASSRWGVRSRTRSGSG GEREPRDRPSERPRLV LPERAFGPRTPRAPRRRRRLLLSPPPRPPPPL DREPRAPGPWLCPSRAGTAQDPARIRERRGR VAGGAAGPAMELRARGWWLLCAAAALVAC
						ARGDPASKSRSCGEVRQIYGAKGFSSS\DVPQ AEISGEHLRICPQGYTCCTSEMEENLANRSHA ELETALRDSSRVLQAMLATQLRSFDDHFQHL LNDSERTLQATFPGAFGELYTQNARAFRDLY SELRLYYRGANLHLEETLAEFWARLLERLFK QLHPQLLLPDDYLDCLGKQAEALRPF\GEAP\ RELRLRAT\RA\FVAAR\SFVQGLGVAS\DVVR KVAQVPLG\PEC\SRAVIEAGSYC\ALHCVGVP GARPCPDYCRNVLKGCLANQADLDAEWRNL
						LDSMVLITDKFWGTSGVESVIGSVHTWLAEA INALQDNRDTLTAKVIQGCGNPKVNPQGPGP EEKRRGKLAPRERPPSGTLEKLVSEAKAQL RDVQDFWISLPGTLCSEKMALSTASDDRCWN GMARGRYLPEVMGDGLANQINNPEVEVDIT KPDMTIRQQIMQLKIMTNRLRSAYNGNDVDF QDASDDGSGSGSGDGCLDDLCGRKVSRKSSS SRTPLTHALPGLSEQEGGKTSAASCPQPPTFL LPLLLFLALTVARPRWR
831	2181	A	6808	2	1522	ASRHGMTPGALLMLLGALGPPLAPGVRGSEA EGRLREKLFSGYDSSVRPAREVGDRVRVSVG LILAQLISLNEKDEEMSTKVYLDLEWTDYRLS WDPAEHDGIDSLRITAESVWLPDVVLLNNND GNFDVALDISVVVSSDGSVRWQPPGIYRSSCS IQVTYFPFDWQNCTMVFSSYSYDSSEVSLQT GLGPDGQGHQEIHIHEGTFIENGQWENIHKPS RLIQPPGDPRGGREGQRQEVIFYLIIRRKPLFY LVNVIAPCILITLLAIFVFYLPPDAGEKMGLSIF ALLTLTVFLLLLADKVPETSLSVPIIKYLMFT MVLVTFSVILSVVVLNLHHRSPHTHQMPLWV RQIFIHKLPLYLRLKRPKPERDLMPEPPHCSSP GSGWGRGTDEYFIRKPPSDFLFPKPNRFQPEL SAPDLRRFIDGPNRAVALLPELREVVSSISYIA RQLQEQEDHDALKEDWQFVAMVVDRLFLW TFIIFTSVGTLVVIFLDATYHLPPPDFFP
832	2182	A	6824	71	1079	ETMAKNPPENCEDCHILNAEAFKSKKICKSLK ICGLVFGILALTLIVLFWGSKHFWPEVPKKAY DMEHTFYSNGEKKKIYMEIDPVTRTEIFRSGN GTDETLEVHDFKNGYTGIYFVGLQKCFIKTQI KVIPEFSEPEEEIDENEEITITFFEQSVIWVPAE KPIENRDFLKNSKILEICDNVTMYWINPTL\IS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Trytophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion GTFAKQLHHNFAPIILVSELQDFEEEGEDLHFP ANEKKGIEQNEQWVVPQVKVEKTRHARQAS EEELPINDYTENGIEFDPMLDERGYCCIYCRR GNRYCRRVCEPLLGYYPYPYCYQGGRVICRV IMPCNWWVARMLGRV
833	2183	A	6846	116	602	EAEGEQVÖGAKCCGDAPHVENREEETARIGP GVMESKEERALNNLIVENVNQENDEKDEKE QVANKGEPLALPLNVSEYCVPRGNRRRFRVR QPILQYRWDIMHRLGEPQARMREENMERIGE EVRQLMEKLREKQLSHSLRAVSTDPPHHDHH DEFC\LMP
834	2184	A	6851	3	2024	PNGVALLHLPGAAVIPNTNYMFQDALGGRSR GSREESPAPSRAPASASLWRRLVVVEAKMAA HAAAAAQAAAAQAAHAEAADSWYLALLGF AEHFRTSSPPKIRLCVHCLQAVFPFKPPQRIEA RTHLQLGSVLYHHTKNSEQARSHLEKAWLIS QQIPQFEDVKFEAASLLSELYCQENSVDAAKP LLRKAIQISQQTPYWHCRLLFQLAQLHTLEKD LVSACDLLGVGAEYARVVGSEYTRALFLLSK GMLLLMERKLQEVHPLLTLCGQIVENWQGN PIQKESLRVFFLVLQVTHYLDAGQVKSVKPC LKQLQQCIQTISTLHDDEILPSNPADLFHWLP KEHMCVLVYLVTVMHSMQAGYLEKAQKYT DKALMQLEKLKMLDCSPILSSFQVILLEHIIM CRLVTGHKATALQEISQVCQLCQQSPRLFSN HAAQLHTLLGLYCVSVNCMDNAEAQFTTAL RLTNHQELWAFIVTNLASVYIREGNRHQEVV LYSLLERINPDHSFPVSSHCLRAAAFYVRGLF SFFQGRYNEAKRFLRETLKMSNAEDLNRLTA CSLVLLGHIFYVLGNHRESNNMVVPAMQLAS KIPDMSVQLWSSALLRDLNKACGNAMDAHE AAQMHQNFSQQLLQDHIEACSLPEHNLITWT DGPPPVQFQAQNGPNTSLASLL
835	2185	A	6855	315	1268	PTRRPILPLTSPKAISVPSPLQGKQHTLVKSCL SVSGIGGFLVSLSSRMKLQTLAVSVTALKFWS AYVPCQTQDRDALRLTLEQIDLIRRMCASYSE LELVTSAKALNDTQKLACLIGVEGGHSLDNS LSILRTFYMLGVRYLTLTHTCNTPWAESSAK GVHSFYNNISGLTDFGEKVVAEMNRLGMMV DLSHVSDAVARRALEVSQAPVIFSHSAARGV CNSARNVPDDILQLLEEERWAFVMVSLFHGE LIQWQPIRPMCSTVADHFDHIKAVIGSKFIGI GGDYDGAGK YRKKTTCKAPWRTSSRMSS PPRSRPSCWRKKVGPGRPWWWGGTGPPGQG
		·				RPEIRLLPLPMTGACGAVAASRTGSSGPG/SSL PNGHGGKGSGLANGLAGNP\GHLGLGSSFGT GPGSGRPPP
837	2187	A	6863	2	1615	VLRGQRGPAGGLAEERRRGRNEWRIHDVTT APFPGLVQRRSRLLIVSQVRYFLKNKVSPDLC NEDGLTALHQCCIDNFEEIVKLLLSHGANVN AKDNELWTPLHAAATCGHINLVKILVQYGA DLLAVNSDGNMPYDLCEDEPTLDVIETCMAY QGITQEKINEMRVAPEQQMIADIHCMIAAGQ DLDWIDAQGATLLHIAGANGYLRAAELLLDH GVRVDVKDWDGWEPLHAAAFWGQMQMAE LLVSHGAN\LNARTSMDEMPIDLCEEEEFKVL LLELK\HKHDVIMKSQLRHKSSLSRRTSHRQA S/SVGKVVRRTQPVGTGPNL\YRKEYE/GEEAI LWQRSA\AEDQRTSTYNGDIRET\RTDQENKD

SEQ ID	SEO ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	1100	in in	nucleotide	1	D=Aspartic Acid, E=Glutamic Acid,
eotide	seq-	1	USSN	1	location	F=Phenylalanine, G=Glycine, H=Histidine,
I.				location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
ł	i	İ	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
ļ	Į.		1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	1		[peptide	-	/=possible nucleotide deletion, \=possible
İ			l	sequence		nucleotide insertion
				·	· · · · · · · · · · · · · · · · · · ·	PNPRLEK\PVLLSEFPTKIPRGELDMPVENGLR
1		l				ADVEAVOVAL ANCOVIDENTIAL VENOUR
			1			APVSAYQYALANGDVWKVHEVPDYSMAYG
			1			NPGVADATPPWSSYKEQSPQTLLELKRQRAA
İ	1					AKLLSHPFLSTHLGSSMARTGESSSEGKAPLI
		İ	ļ			GGRTSPYSSNGTSVYYTVTSGDPPLLKFKAPI
						EEMEEKVHGCCRIS
838	2188	Α -	6865	6291	739	AGPLEPRVQGAMALQLWALTLLGLLGAGAS
ł	ĺ		i l		•	LRPRKLDFFRSEKELNHLAVDEASGVVYLGA
- 1	İ					VNALYQLDAKLQLEQQVATGPVLDNKKCTP
						PIEASQCHEAEMTDNVNQLLLVDPPRKRLVE
1						CCOLLYCICAL DALCATEL DI EXEDOCORNO
l						CGQLLKGI\CALRALSNISLRLFYEDGSGEKSF
1		j] ;			VASNDEGVATVGLVSSTGPGGDRVLFVGKG
		1				NGPHDNGIIVSTRLLDRTDSREAFEAYTDHAT
1						YKAGYLSTNTQQFVAAFEDGPYVFFVFNQQD
	1					KHPARNRTLLARMCREDPNYYSYLEMDLQC
-	'					RDPDIHAAAFGTCLAASVAAPGSGRVLYAVF
i			į			SRDSRSSGGPGAGLCLFPLDEVHAKMEANRN
1	1		i i		,	ACYTGTREARDIFYKPFHGDIQCGGHAPGSSK
						SFPCGSEHLPYPLGSRDGLRGTAVLORGGLN
						LTAVTVAAENNHTVAFLGTSDGRILKVYLTP
						DGTSSEYDSILVEINKRVKRDLVLSGDLGSLY
j i]	}		
						AMTQDKVFRLPVQECLSYPTCTQCRDSQDPY
						CGWCVVEGRCTRKAECPRAEEASHWLWSRS
						KSCVAVTSAQPQNMSRRAQGEVQLTVSPLPA
						LSEEDELLCLFGESPPHPARVEGEAVICNSPSS
						IPVTPPGQDHVAVTIQLLLRRGNIFLTSYQYPF
					;	YDCRQAMSLEENLPCISCVSNRWTCQWDLR
					1	YHECREASPNPEDGIVRAHMEDSCPQFLGPSP
	l i		1		. 1	LVIPMNHETDVNFQGKNLDTVKGSSLHVGSD
			1			LLKFMEPVTMQESGTFAFRTPKLSHDANETL
			j			PLHLYVKSYGKNIDSKLHVTLYDCSFGRSDC
						SLCRAANPDYRCAWCGGQSRCVYEALCNTT
	1	ł	1	ł		SECPPPVITRIQPETGPLGGGIRITILGSNLGVQ
		i	ı	1		AGDIQRISVAGRNCSFQPERYSVSTRIVCVIEA
						A ETRET COVERNO STORY CROPPED OF
1					ĺ	AETPFTGGVEVDVFGKLGRSPPNVQFTFQQP
[ł			KPLSVEPQQGPQAGGTTLTIHGTHLDTGSQED
		1				VRVTLNGVPCKVTKFGAQLQCVTGPQATRG
[[Í		ł		QMLLEVSYGGSPVPNPGIFFTYRENPVLRAFE
		İ	}	!	.	PLRSFASGGRSINVTGQGFSLIQRFAMVVIAEP
		J		1	;	LQSWQPPREAESLQPMTVVGTDYVFHNDTK
		ľ				VVFLSPAVPEEPEAYNLTVLIEMDGHRALLRT
		1		ļ	1	EAGAFEYVPDPTFENFTGGVKKQVNKLIRAR
		l		ĺ	ŀ	GTNLNKAMTLQEAEAFVGAERCTMKTLTET
1 1		- 1	İ	1	ľ	DLYCEPPEVQPPKRRQKRDTTHNLPEFIVKF
1		l				GCDEWAI COMEADAMANDI DI COME COME CONTROL COME CONTROL COME CONTROL COME CONTROL COME COME COME COME COME COME COME COME
		ĺ	İ			GSREWVLGRVEYDTRVSDVPLSLILPLVIVPM
		ļ	1		٠ ,	VVVIAVSVYCYWRKSQQAEREYEKIKSQLEG
1		1		1	· · · · · · · · · · · · · · · · · · ·	LEESVRDRCKKEFTDLMIEMEDQTNDVHEAG
		- 1		ļ		IPVLDYKTYTDRVFFLPSKDGDKDVMITGKL
				į		DIPEPRRPVVEQALYQFSNLLNSKSFLINFIHT
]	-	İ	-	ļ	L'ENQPEFSARAKVYFASLLTVALHGKLEYYT
1 1	j	1				DIMHTLFLELLEQYVVAKNPKLMLRRSETVV
	ļ	ĺ	į	1	1	ERMLSNWMSICLYQYLKDSAGEPLYKLFKAI
1 1		ł	ļ	ŀ		KHQVEKGPVDAVQKKAKYTLNDTGLLGDD
1		ŀ	- 1			VEVADI TUQUIVODE CUDA IDIRIZIRA I ICONTO
1 1			1]		VEYAPLTVSVIVQDEGVDAIPVKVLNCDTISQ
j		i		ŀ	l	VKEKIIDQVYRGQPCSCWPRPDSVVLEWRPG
1		l		ļ		STAQILSDLDLTSQREGRWKRVNTLMHYNVR
j i			ľ			DGATLILSKVGVSQQPEDSQQDLPGERHALL
į į		1		· 1	į.	EEENRVWHLVRPTDEVDEGKSKRGSVKEKE
}	ļ	1		į	ľ	RTKAITEIYLTRLLSVKGTLQQFVDNFFOSVL
		J				APGHAVPPAVKYFFDFLDEQAEKHNIQDEDTI

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	1		
nucl-	peptide	HOG	in in	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
	1 * *			nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	l			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
i				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1		1	ł	peptide	1	/=possible nucleotide deletion, \=possible
ļ				sequence		nucleotide insertion
						HIWKTNSLPLRFWVNILKNPHFIFDVHVHEVV
ļ]	}				DASLSVIAQTFMDACTRTEHKLSRDSPSNKLL
1	ĺ	1	Ì		1	YAKEISTYKKMVEDYYKGIRQMVQVSDQDM
						NTHLAEISRAHTDSLNTLVALHQLYQYTQKY
ļ						YDEINALEEDPAAQKMQLAFRLQQIAAALE
1						NKVTDL
839	2189	A	6872	ļ. 	1485	
039	2109	A	00/2	1	1465	RARRLALQCHVCVCALTPGEQSGRRLPGQT
	ŀ	1	i			WLMFSCFCFSLQDNSFSSTTVTECDEDPVSLH
]				j	EDQTDCSSLRDENNKENYPDAGALVEEHAPP
			Í			SWEPQQQNVEATVLVDSVLRPSMGNFKSRKP
						KSIFKAESGRSHGESQETEHVVSSQSECQVRA
1	1		}			GTPAHESPQNNAFKCQET\VRL\QPRIDQRTAT
j i	}	1]			SPKDAFETR\QDLNEEEAAQVHGVKDPAPAS
	İ	İ				TQSVLA\DGTDSADPSPVHKDGQNEADSAPE
						DLHSVGTSRLLL/YHITDGDNPTAVRHGCSL/F
1	1				}	SGOSORFNLDPESAPSPPSTOOFMMPRSSSRC
						SCGDGKEPQTITQLTKHIQSLKRKIRKFEEKFE
1	ļ	İ				QEKKYRPSHGDKTSNPEVLKWMNDLAKGRK
1	ľ	ĺ	1			
	i	ŀ				QLKELKLKLSEEQGSAPKGPPRNLLCEQPTVP
	i	ĺ				RENGKPEAAGPEPSSSGEETPDAALTCLKERR
1						EQLPPQEDSKVTKQDKNLIKPLYDRYRIIKQIL
						STPSLIPTIVSQDTCMLLLCTDV
840	2190	A	6873	2	2054	FFRFYFSFIRLFAMSLADLTKTNIDEHFFGVAL
						ENNRRSAACKRSPGTGDFSRNSNASNKSVDY
						SRSQCSCGSLSSQYDYSEDFLCDCSEKAINRN
			[YLKQPVVKEKEKKKYNVSKISQSKGQKEISV
						EKKHTWNASLFNSQIHMIAQRRDAMAHRILS
						ARLHKIKGLKNELADMHHKLEAILTENOFLK
1 1		l				QLQLRHLKAIGKYENSQNNLPQIMAKHQNEV
						KNLRQLLRKSQEKERTLSRKLRETDSQLLKT
1						
1						KDILQALQKLSEDKNLAEREELTHKLSIITTK
1 1						MDANDKKIQSLEKQLRLNCRAFSRQLAIETR
1						KTLAAQTATKTLQVEVKHLQQKLKEKDREL
						EIKNIYSHRILKNLHDTEDYPKVSSTKSVQAD
!						RKILPFTSMRHQGTQKSDVPPL/TTKGKKATG
1 1						NIDHKEKSTEINHEIPHCVNKLPKQEDSKRKY
						EDLSGEEKHLEVQILLENTGRQKDKKEDQEK
'						KNIFVKEEQELPPKIIEVIHPERESNQEDVLVR
				ļ		EKFKRSMQRNGVDDT\LGKGTAPYTKGPLRQ
]]				ļ	· ·	RRHYSFTEATENLHHGLPASGGPANAGNMR
					ļ	YSHSTGKHLSNREEMELEHS\DSGYEPSFGKS
						SRIKVKDTTFRDKKSSLMEELFGSGYVLKTD
						QSSPGVAKGSEEPLOSKESHPLPPSOASTSHA
	,					FGDSKVTVVNSIKPSSPTEGKRKIII
841	2191	Α	6874	3	2867	SSRTREMEEKEILRRQIRLLQGLIDDYKTLHG
		**	30,1	~	230,	NAPAPGTPAASGWQPPTYHSGRAFSARYPRP
						SRRGYSSHHGPSWRKKYSLVNRPPGPSDPPA
1						DHAVRPLHGARGGQPPVPQQHVLERQVQLS
			ļ			QGQNVVIKVKPPSKSGSASASGAQRGSLEEFE
						DTPWSDQRPREGEGEPPRGQLQPSRPTRARG
						TCSVEDPLLVCQKEPGKPRMVKSVGSVGDSP
	1		ľ	1		REPRRTVSESVIAVKASFPSSALPPRTGVALG
			ļ			RKLGSHSVASCAPQLLGDRRVDAGHTDQPVP
1		ļ	1			SGSVGGPARPASGPRQAREASLVVTCRTNKF
]		Ì	.	1		RKNNYKWVAASSKSPRVARRALSPRVAAEN
[ĺ	ĺ	Í	1		VCKASAGMANKVEKPQLIADPEPKPRKPATS
	ļ			ļ		SKPGSAPSKYKWKASSPSASSSSFRWQSEAG
]			İ		SKDHASQLSPVLSRSPSGD\RPALAHSGLKPLS
1 1			i	1		PARTICIPATION OF THE PROPERTY
	1	ı	ı	ı	1	GETPI SAVKVKTDTKIIDDD GOTOI DGDVVOG
	Ī					GETPLSAYKVKTRTKIIRRRGSTSLPGDKKSG TSPAATAKSHLSLRRRQALRGKSSPVLKKTPN

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methioriine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
					·	KGLVQVTKHRLCRLPPSRAHLPTKEASSLHA VRTAPTSKVIKTRYRIVKKTPASPLSAPPFPLS LPSWRARRLSLSRSLVLNRLRPVASGGGKAQ PGSPWWRSKGYRCIGGVLYKVSANKLSKTSG QPSDAGSRPLLRTGRLDPAGSCSRSLASRAVQ RSLAIIRQARQRREKRKEYCMYYNRFGRCNR GERCPYIHDPEKVAVCTRFVRGTCKKTDGTC PFSHHVSKEKMPVCSYFLKGICSNSNCPYSHV YVSRKAEVCSDFLKGYCPLGAKCKKHTLLC PDFARRGACPRGAQCQLLHRTQKRHSRRAAT SPAPGPSDATARSRVSASHGPRKPSASQRPTR QTPSSAALTAAAVAAPPHCPGGSASPSSSKAS SSSSSSSSPPASLDHE\APSLQEAALAAACSNR LCKLPSFISLQSSPSPGAQPRVRAPRAPLTKDS GKPLHIKPRL
842	2192	A	6898	506	2071	WPDLVHTWSSEEAMGSCCSCPDKDTVPDNH RNKFKVINVDDDGNELGSGIMELTDTELILYT RKRDSVKWHYLCLRRYGYDSNLFSFESGRRC QTGQGIFAFKCARAEELFNMLQEIMQNNSIN VVEEPVVERNNHQTELEVPRTPRTPTTPGFAA QNLPNGYPRYPSFGDASSHPSSRHPSVGSARL PSVGEESTHPLLVAEEQVHTYVNTTGVQEER KNRTSVHVPLEARVSNAESSTPKEEPSSIEDR DPQILLEPEGVKFVLGPTPVQKQLMEKEKLE QLGRDQVSGSGANNTEWDTGYDSDERRDAP SVNKLVYENINGLSIPSASGVRRGRLTSTSTSD TQNINNSAQRRTALLNYENLPSLPPVWEARK LSRDEDDNLGPKTPSLNGYHNNLDPMHNYV NTENVTVPASAHKIEYSRRRDCTPTVFNFDIR RPSLEHRQLNYIQVDLEGGSDSDNPQTPKTPT TPLPQTPTRRTELYAVIDIERTAAMSNLQKAL
843	2193	A	6919	2	663	PRDDGTSR/KTRHNST/DLPL AGRPGTTHASGKMAYQSLRLEYLQIPPVSRA YTTACVLTTAAVQLELITPFQLYFNPELIFKHF QIWRLITNFLFFGPVGFNFLFNMIFLYRYCRM LEEGSFRGRTADFVFMFLFGGFLMTLFGLFVS L/VFLGPGLYNN/GSSMCGAE\EPLCPHELLRP SQLPGPLSALGAHGIFLVVGELNHCGPFGYCS WTHIFFLGRCISQSTWWNKNSENTTYFESYF
844	2194	A	6928	902	366	HRLCMPIQGACGERME/FSLLLPGLECNGVIL AHCNLRLPGSSNSPASASQVAGITGVCHHAR LIFVFSVETGFLHAGQAGLELLTSGDPPASAS QSAGITGKSQHTRPGYEFIIPYSAAQEDALKA LM
845	2195		6939	1660	317	LYPENLGESLFPILLLPPPWPDGGRPCCVEMS TRAKKLRIWRILEEKESVAGAVQTLLLRSQE GGVITSAAASTLSEPPRRTQESRTRTRALGLPT LPMEKLAASTESPPRRTQESRTRTRALGLPT LPMEKLAASTEPQGPRPVLGRESVQVPDDQD FRSFRSECEAEVGWNLTYSRAGVSVWVQAV EMDRTLHKIKCRMECCDVPAETLYDVLHDIE YRKKWDSNVIETFDIARLTVNADVGYYSWR CPKPLKNRDVITLRSWLPMGADYIIMNYSVK HPKYPPRKDLVRAVSIQTGYLIQSTGPKSCVIT YLAQVDPKGSLPKWVVNKSSQFLAPKAMKK MYKACLKYPEWKQKHL\PHFKPWL\HPEQSP LPSLALS\ELSVQHADS\LENIDESAV\AESREE R\MGGAGGEG\SDDDTSLYAEAPHRFRETETG PGAGRALGAAAAPALSPLHPPGTWWHRARP RRVLQPGWTEPQ

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
846	2196	Α	6944	42	2672	RRKMAGCRGSLCCCCRWCCCCGERETRTPE ELTILGETQEEEDEILPRKDYESLDYDRCINDP YLEVLETMDNKKGRRYEAVK WMVVFAIGV CTGLVGLFVDFFVRLFTQLKFGVVQTSVEECS QKGCLALSLLELLGFNLTFFVLESLLGLIEPVE AGSGITEGKCYLYARQVPGLVRLPTLLWKAL GVLLTVAAMLLI\GLGSPMIHSGSVVGAGLPQ FQSISLRKIQFNFPYFRSDRYGK\DKRDFVSAG AAAGVAAAFGAPIGGTLFSLEEGSSFWNQGL TWKVLFCSMSATFTLNFFRSGIQFGSWGSFQL PGLLNFGEFKCSDSDKKCHLWTAMDLGFFV VMGVIGGLLGATFNCLNKRLAKYRMRNVHP KPKLVRVLESLLVSLVTTVVVFVASMVLGEC RQMSSSSQIGNDSFQLQVTEDVNSSIKTFFCP NDTYNDMATLFFNPQESAILQLFHQDGTFSPV TLALFFVLYFLLACWTYGISVPSGLFVPSLLC GAAFGRLVANVLKSYIGLGHIYSGTFALIGAA AFLGGVVRMTISLTVILIESTNEITYGLPIMVT LMVGKWTGDFFNKGI\YDIHVGLRGVPLLEW ETEVEMDKLRASDIMEPNLTYVYPHTRIQSLV SILRTTVHHAFPVVTENRGNEKEFMKGNQLIS NNIKFKKSSILTRAGEQRKRSQSMKSYPSSEL RNMCDEHIASEEPAEKEDLLQQMLERRYTPY PNLYPDQSPSEDWTMEERFRPLTFHGLILRSQ LVTLLVRGVCYSESQSSASQPRLSYAEMAED YPRYPDIHDLDLTLLNPRMIVDVTPYMNPSPF TVSPNTHVSQVFNLFRTMGLRHLPVVNAVGE IVGITTRHNLTYEFLQARLRQHYQTI
847	2197	A	6951		1994	NTNSSSVTNSAAGVEDLNIVQVTVPDNEKER LSSIEKIKQLREQVNDLFSRKFGEAIGVDFPVK VPYRKITFNPGCVVIDGMPPGVVFKAPGYLEI SSMRRILEAAEFIKFTVIRPLPGLELSNGEYST VGKRKIDQEGRVFQEKWERAYFFVEVQNIST CLICKRSMSVSKEYNLRHYQTNHSKHYDQY MERMRDEKLHELKKGLRKYLLGLSDTECPE QKQVFANPSPTQKSPVQPVEDLAGNLWEKLR EKIRSFVAYSIAIDEITDINNTTQLAIFIRGVDE NFDVSEELLDTVPMTGTKSGNEIFSRVEKSLK NFCINWSKLVSVASTGTPPMVDANNGLVTKL KSRVATFCKGAELKSICCIIHPESLCAQ\KLKM DHVMDVVVKSVNWICSRGLNHSEFTTLLYEL DSQYGSLLYYTEIKWLSRGLVLKRFFESLEEI DSFMSSRGKPLPQLSSIDWIRDLAFLVDMTM HLNALNISLQGHSQIVTQMYDLIRAFLAKLCL WETHLTRNNLAHFPTLKLVSRNESDGLNYIP KIAELKTEFQKRLSDFKLYESELTLFSSPFSTKI DSVHEELQMEVIDLQCNTVLKTKYDKVGIPE FYKYLWGSYPKYKHHCAKILSMFGSTYICEQ LFSIMKLSKTKYCSQLKDSQWDSVLHIAT
848	2198	A	6985	3	289	SVQYLPGRPTRTHASTDAPLMLKFTPLPSKTK ASAPVQCLLLMAATFSPQGLAKPHSGTIPIT\C CFNAINTKIPIQRLESYTRITNIQCPKEAVM
849	2199	A	6999	963	5	LDFLCHRDMGDNITSITEFLLLGFPVGPRIQM LLFGLFSLFYVFTLLGNGTILGLISLDSRLHAP MYFFLSHL\AVVDIA YACNTVPRMLVNLLHP AKPISFAGRMMQTFLFSTFAVTECLLLVVMS YDLYV\AICHPLRYLAIMTWRVCITLAVTSWT TGVLLSLIHLVLLLPLPFCRPQKIYHFFCEILA VLKLACADTHINENMVLAGAISGLVGPLSTIV VSYMCILCAILQIQSREVQRKAFCTCFSHLCVI

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid. E=Glutamic Acid.
nucl-	peptide	"""	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
			***	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		1		peptide	sequence	1=1 yrosine, A=Unknown, *=Stop codon,
	1	1		1		/=possible nucleotide deletion, \=possible
<u> </u>	 	ļ		sequence		nucleotide insertion
	1					GLFYGTAIIMYVGPRYGNPKEQKKYLLLFHS
950	2200	<u> </u>	7001			LFNPMLNPLICSLRNSEVKNTLKRVLGVERAL
850	2200	Α	7001	1	1011	MGNDSVSYEYGDYSDLSDRPVDCLDGACLAI
i						DPLRVAPLPLYAAIFLVGVPGNAMVAWVAG
		İ				KVARRRVGATWLLHLAVADLLCCLSLPILAV
	}					PIARGGHWPYGAVGCRALPSIILLTMYASVLL
		Ì				LAALSADLCFLALGPAW\CLRFS/GACGVQVA
}		i				CGAAWTLALLLTVPSAIYRRLHQEHFPARLQ
1		ĺ				CVVDYGGSSSTENAVTAIRFLFGFLGPLVAVA
						SCHSALLCWAARRCRPLGTAIVVGFFVCWAP
	ľ	1				YHLLGLVLTVAAPNSALLARALRAEPLIVGL
						ALAHSCLNPMLFLYFGRAQLRRSLPAACHW
1	}	1				ALRESQGQDESVDSKKSTSHDLVSEMEV
851	2201	Α	7011	1	2310	AAASPLRMSRKGPRAEVCADCSAPDPGWASI
1			' ' '		20.0	SRGVLVCDECCSVHRSLGRHISIVKHLRHSA
		ł				WPPTLLQMVHTLASNGANSIWEHSLLDPAQV
						OCCUAL KOTOKO KANDIWEED AKKO A AF
						QSGPALKQTPKDKV\HPIKSEFIRAKYQMLAF
						VHKLPCRDDDGVTAKDLSKQLHSSVRTGNLE
ļ	1					TCLRLLSLGAQANFFHPEKGTTPLHVAAKAG
}						QTLQAELLVVYGADPGSPDVNGRTPIDYARQ
						AGHHELAERLVECQYELTDRLAFYLCGRKPD
1						HKNGHYIIPQMADSLDLSELAKAAKKKLQAL
						SNRLFEELAMDVYDEVDRRENDAVWLATQN
						HSTLVTERSAVPFLPVNPEYSATRNQGRQKL
1 1						ARFNAREFATLIIDILSEAKRRQQGKSLSSPTD
1		:			,	NLELSLRSQSDLDDQHDYDSVASDEDTDQEP
						LRSTGATRSNRARSMDSSDLSDGAVTLQEYL
						ELKKALATSEAKVQQLMKVNSSLSDELRRLQ
1						REIHKLQAENLQLRQPPGPVPTPPLPSERAEH
1 1				1		TPMAPGGSTHRRDRQAFSMYEPGSALKPFGG
	•					PPGDELTTRLQPFHSTELEDDAIYSVHVPAGL
						YKIRKGVSASAVPFTPSSPLLSCSQEGSRHTSK
i i				ļ		LSRHGSGADSDYENTQSGDPLLGLEGKRFLE
i						LGKEEDFHPELESLDGDLDPGLPSTEDVILKT
1 1				Ì		EQVTKNIQELLRAAQEFKHDSFVPCSEKIHLA
			}		ļ	VTEMASLFPKRPALEPVRSSLRLLNASAYRLQ
					į	SECRKTVPPEPGAPVDFQLLTQQVIQCAYDIA
<u> </u>			ł	ł		KAAKQLVTITTREKKQ
852	2202	A	7016	484	1777	RISKIQVYYSTGYSSRKMNPTLGLAIFLAVLL
						TVKGLLKPSFSPRNYKALSEVQGWKQRMAA
]			1	1		KELARQNMDLGFKLLKKLAFYNPGRNIFLSP
1 1	j		}	ļ		LSISTAFSMLCLGAQDSTLDEIKQGFNFRKMP
1 1		l	- 1	ľ		EKDLHEGFHYIIHELTQKTQDLKLSIGNTLFID
			l	l		
•			l	ļ	İ	QRLQPQRKFLEDAKNFYSAETILTNFQNLEM
		- 1	ļ	- 1	J	AQKQINDFI/ESKTHGKINNLIENIDPGTVMLL
		[}	l	Ī	ANYIFFRARWKHEFDPNVTKEEDFFLEKNSS
1		,		ļ		VKVPMMFRSGIYQVGYDDKLSCTILEIPYQK
[]		l		Ì		NITAIFILPDEGKLKHLEKGLQVDTFSRWKTL
1		ļ	İ	[ŀ	LSRRVVDVSVPRLHMTGTFDLKKTLSYIGVS
1 1	!	I	-	ļ)	KIFEEHGDLTKIAPHRSLKVGEAVNKAELKM
1	1	l		ļ	l	DERGTEGAAGTGAQTLPMETPLVVKIDKPYL
0.50	-0000					LLIYSEKIPSVLFLGKIVNPIGK
853	2203	A	7017	1	3293	MTHACNPSTLGGQGRRITRSHGRRRSSRGPV
		į	j		ĺ	ARHVAAGAGHENKHGGSRRFPAGVAPRRAM
1	1	1	1		1	ANVSKKVSWSGRDRDDEEAAPLLRRTARPG
1	J	l				GGTPLLNGAGPGAARQSPRSALFRVGHMSSV
1 1	ŀ		J		1	ELDDELLEP\DMDPPHPFPKEIPHNEKLLSLKY
	l	- 1	İ		Į.	ESLDYDNSENQLFLEEERRINHTAFRTVEIKR
				l		WVICALIGILTGLVACFIDIVVENLAGLKYRVI
[ĺ	1	1	ŀ	ł	KGSILPNIDKFTEKGGLSFSLLLWATLNAAFV

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LVGSVIVAFIEPVAAGSGIPQIKCFLNGVKIPH VVRLKTLVIKVSGVILSVVGGLAVGKEGPMI
						HSGSVIAAGISQGRSTSLKRDFKIFEYFRRDTE KRDFVSAGAAAGVSAAFGAPVGGVLFSLEEG ASFWNQFLTWRIFFASMISTFTLNFVLSIYHG NMWDLSSPGLINFGRFDSEKMAYTIHEIPVFI AMGVVGGVLGAVFNALNYWLTMFRIRYIHR PCLQVIEAVLVAAVTATVAFVLIYSSRDCQPL QGGSMSYPLQLFCADGEYNSMAAAFFNTPEK SVVSLFHDPPGSYNPLTLGLFTLVYFFLACWT YGLTVSAGVFIPSLLIGAAWGRLFGISLSYLTG AAIWADPGKYALMGAAAQLGGIVRMTLSLT VIMMEATSNVTYGFPIMLVLMTAKIVGDVFIE GLYDMHIQLQSVPFLHWEAPVTSHSLTAREV MSTPVTCLRRREKVGVIVDVLSDTASNHNGF PVVEHADDTQPARLQGLILRSQLIVLLKHKVF VERSNLGLVQRRLRLKDFRDAYPRFPPIQSIH VSQDERECTMDLSEFMNPSPYTVPQEASLPR VFKLFRALGLRHLVVVDNRNQVVGLVTRKD LARYRLGKRGLEELSLAQTGPKAQATAEGRV AGAAQQPCQLRAVTLEDLGLLLAGGLASPEP LSLEELSERYESSHPTSTASVPEQDTAKHWNQ LEQWVVELQAEVACLREHKQRCERATRSLL RELLQVRARVQLQGSELRQLQQEARPAAQAP EKEAPEFSGLQNQMQALDKRLVEVREALTRL RRRQVQQEAERRGAEQEAGLRLAKLTDLLQ QEEQGREVACGALQKNQEDSSRRVDLEVAR
854	2204	A	7037	139		AGTWEPRPYDQAKETGAPGSQPPVPPMELRP WLLWVVAATGTLVLLAADAQGQKVFTNTW AVRIPGGPAVANSVARKHGFLNLGQIFGDYY HFWHRGVTKRSLSPHRPRHSRLQREPQVQWL EQQVAKRRTKRDVYQEPTDPKFPQQWYL\SG VTQ\RDLMVKAAWAQGYTGHGIVVSILDDGI EKNHPDLAGNYDPGASFDVNDQDPDPQPRY TQMNDNRHGTRCAGEVAAVANNGVCGVGV AYNARIGGVRMLDGEVTDAVEARSLGLNPN HIHIYSASWGPEDDGKTVDGPARLAEEAFFR GVSQGRGGLGSIFVWASGNGGREHDSCNCD GYTNSIYTLSISSATQFGNVPWYSEACSSTLA TTYSSGNQNEKQIVITDLRQKCTESHTGTSAS APLAAGIIALTLEANKNLTWRDMQHLVVQTS KPAHLNANDWATNGVGRKVSHSYGYGLLD AGAMVALAQNWTTVAPQRKCIDILTEPKDI GKRLEVRKTVTACLGEPNHITRLEHAQARLT LSYNRRGDLAIHLVSPMGTRSTLLAARPHDY SADGFNDWAFMTTHSWDEDPSGEWVLEIEN TSEANNYGTLTKFTLVLYGTAPEGLPVPPESS GCKTLTSSQACVVCEGFSLHQKSCVQHCPP GFAPQVLDTHYSTENDVETIRASVCAPCHAS CATCQGPALTDCLSCPSHASLDPVEQTCSRQS QSSRESPPQQQPPRLPPEVEAGQRLRAGLLPS HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGFS FRGVKVYTMDRGLISYKGLPPEAWQEECPSD
855	2205	A	7058	3	1441	SEEDEGRGERTAFIKDQSAL QRPASQLLAPFAAEALPGAPRAAMAQHFSLA ACDVVGFDLDHTLCRYNLPESAPLIYNSFAQF LVKEKGYDKELLNVTPEDWDFCCKGLALDL EDGNFLKLANNGTVLRASHGTKMMTPEVLA EAYGKKEWKHFLSDTGMACRSGKYYFYDN

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion YFDLPGALLCARVVDYLTKLNNGQKTFDFW KDIVAAIQHNYKMSAFKENCGIYFPEIKRDPG RYLHSRPESVKKWLRQLKNAGKILLLITSSHS DYCRLLCA\YILGNDFTDLFDIVITNALKPGFF SHLPSQRPFRTLENDEEQEALPSLDKPGWYSQ GNAVHLYELLKKMTGKPEPKVVYFGDSMHS DIFPARHYSNWETVLILEELRGDEGTRSORPE
					;	ESEPLEKKGKYEGPKAKPLNTSSKKWGSFFN DSVLGLENTEDSLVYTWSCKRISTYSTIAIPSI EAIAELPLDYKFTRFSSSNSKTAGYYPNPPLV LSSDETLISK
856	2206	A	7082	396	1635	SSPSVFEFEHAVQPVFTMEFLKTCVLRRNACT AVCFWRSKVVQKPSVRRISTTSPRSTVMPAW VIDKYGKNEVLRFTQNMMMPIHYPNEVIVK VHAASVNPIDVNMRSGYGATALNMKRDPLH VKIKGEEFPLTLGRDVSGVVMECGLDVKYFK PGDEVWAAVPPWKQGTLSEFVVVSGNEVSH KPKSLTHTQAASLPYVALTAWSAINKVGGLN DKNCTGKRVLILGASGGVGTFAIQVMKAWD AHVTAVCSQDASELVRKLGADDVIDYKSGSV EEQLKSLKPFDFILDNVGGSTETWAPDFLKK WSGATYVTLVTPFLLNMDRLGIADGMLQTG VTVGSKALKHFWKGVHYRWAFFMASGPCL DDIAELVDAGKIRPVIEQTFPFSKVPEAFLKV ERGHARGKTVINVV
857	2207	A	7088	320	2417	LRRRKMTPQSLLQTTLFLLSLLFLVQGAHGR GHREDFRFCSQRNQTHRSSLHYKPTPDLRISIE NSEEALTVHAPFPAAHPASRSFPDPRGLYHFC LYWNRHAGRLHLLYGKRDFLLSDKASSLLCF QHQEESLAQGPPLLATSVTSWWSPQNISLPSA
			•		· · · ·	ASFTFSFHSPPHTGAHNASVDMCELKRDLQL LSQFLKHPQKASRRPSAAPASQQLQSLESKLT SVRFMGDMGSFEEDRINATVWKLQPTAGLQ DLHIHSRQEEEQSEIMEYSVLLPRTLFQRTKG RSGEAEKRLLLVDFSSQALFQDKNSSQVLGE KVLGIVVQNTKVANLTEPVVLTFQHQLQPKN VTLQCVFWVEDPTLSSPGHWSSAGCETVRRE TQTSCFCNHLTYFAVLMVSSVEVDAVHKHY LSLLSYVGCVVSALACLVTIAAYLCSRVPLPC RRKPRDYTIKVHMNLLLAVFLLDTSFLLSEPV ALTGSEAGCRASAIFLHFSLLTCLSWMGLEG YNLYRLVVEVFGTYVPGYLLKLSAMGWGFPI FLVTLVALVDVDNYGPILLAVHRTPEGVIYPS MCWIRDSLVSYITNLGLFSLVFLFNMAMLAT MVVQILRLRPHTQKWSHVLTLLCLSLVLGILP WALIFFSFASGTFQLVVLYLFSITTSFQGFLIFI WYWSMRLQARGGPSPLKSNSDSARLPISSGS TSSSRI
858	2208	A	7091	185	415	DAGAVKSSDTNIWFRGMCDDKKGHRCPS*G QPQHFHVAFHTEAEGAMFYFRLHVIHRVMQS QQQLFPSTLFSWLLE
859	2209	Α	7136	3	302	FFFWRQSLALLPRLECSGATGAHCNLHFPGSS DCPTSAS*IAGITGACYHAWLLFVFLAETGFH HVGQGGLELLTSSDPSGSASQSAGITGVSHCT WPI
860	2210	A	7156	23	591	ALSTETRTPDMRRLLLVTSLVVVLLWEAGAV PAPKVPIKMQVKHWPSEQDPEKAWGARVVE PPEKDDQLVVLFPVQKPKLLTTEEKPRGQGR GPILPGTKAWMETEDTLGRVLSPEPDHDSLY

NO: of nucleotide eotide sequence were leading to last amino acid residue of peptide sequence leading nucleotide location leading nucleotide location location leading nucleotide location location leading nucleotide location location leading nucleotide lo	istidine, >, Proline, 1e,
eotide sequence USSN 09/496 09/496 orresponding uence uence uence USSN 09/496 orresponding to last amino acid residue of peptide sequence sequence sequence uence USSN 09/496 orresponding to last amino acid residue of peptide sequence sequence sequence sequence uence USSN 09/496 orresponding to last amino acid residue of peptide sequence sequence sequence sequence uence USSN 09/496 orresponding to last amino acid residue of peptide sequence sequence yellow 15-corresponding to last amino acid residue of peptide sequence yellow 15-corresponding to last amino acid residue of peptide sequence yellow 15-corresponding to last amino acid residue of peptide sequence yellow 15-corresponding to last amino acid residue of peptide sequence yellow 15-corresponding to last amino acid residue of peptide sequence yellow 15-corresponding to last amino acid residue of peptide sequence yellow 15-corresponding to last amino acid residue of peptide sequence yellow 15-corresponding to last amino acid residue of peptide sequence yellow 15-corresponding to last amino acid residue of peptide sequence yellow 15-corresponding to last amino acid residue of peptide sequence yellow 15-corresponding to last amino acid residue of peptide yellow 15-corresponding to last amino acid residue of peptide yellow 15-corresponding to last amino acid residue of peptide yellow 15-corresponding to last amino acid residue of peptide yellow 15-corresponding to last amino acid residue of peptide yellow 15-corresponding to last amino acid residue of peptide yellow 15-corresponding to last amino acid residue of peptide yellow 15-corresponding to last amino acid residue of peptide yellow 15-corresponding to last amino acid residue of peptide yellow 15-corresponding to last amino acid residue of peptide yellow 15-corresponding to last amino acid residue of peptide yellow 15-corresponding to last amino acid residue of peptide yellow 15-corresponding to last amino acid residue of peptide yellow 15-corresponding to last amino acid residue of peptide yellow 15-corre	e, Proline, ne,
sequence uence uence 09/496 correspondi ng to first amino acid residue of peptide sequence sequence sequence of peptide sequence sequence of peptide sequence sequence sequence of peptide sequence sequence of peptide sequence sequence of peptide sequence sequence of peptide sequence of peptide sequence sequence of peptide of peptide sequence of peptide of peptide sequence of peptide o	Proline, ne,
uence 914 ng to first amino acid residue of peptide sequence 914 ng to first amino acid residue of peptide sequence 92=Glutamine, R=Arginine, S=Serin T=Threonine, V=Valine, W=Trypto Y=Tyrosine, X=Unknown, *=Stop /=possible nucleotide deletion, \=po nucleotide insertion HPPPEEDQGEERPRLWVMPNHO DHIYHPQ*GSRGHHCPRPVPRPR CPS	1e,
amino acid residue of peptide sequence Y=Treonine, V=Valine, W=Trypto Y=Tyrosine, X=Unknown, *=Stop of peptide sequence sequence sequence sequence sequence sequence HPPPEEDQGEERPRLWVMPNHO DHIYHPQ*GSRGHHCPRPVPRPF CPS	nhan
peptide /=possible nucleotide deletion, \=po nucleotide insertion HPPPEEDQGEERPRLWVMPNHC DHIYHPQ*GSRGHHCPRPVPRPF CPS	
sequence nucleotide insertion HPPPEEDQGEERPRLWVMPNHC DHIYHPQ*GSRGHHCPRPVPRPF CPS	
HPPPEEDQGEERPRLWVMPNHC DHIYHPQ*GSRGHHCPRPVPRPF CPS	ssible
DHIYHPQ*GSRGHHCPRPVPRPF CPS	
CPS	
	KLLULUFSLF
861 2211 A 7161 1220 1003 NYVCTIAF*EKKMGF*LSLSCLV	/LLFVLFLDCI
LTTTTRIMFHCTYLFASVCLSLL	
KSAMILQ .	
862 2212 A 7211 665 847 LKYYHITMGIYKTGKKVIL*KSS	
YKNIQKLSFSNYVYHQNYVFSS	
863 2213 A 7212 924 1273 HGSSCALGDLAPG*LPSGPVLSS	
LVWDSPSCLPATGPT*GLVLVL0 RGQHEHKRMRAP*SCRVTVNL	
CIKPNYQSPPKECDYNILANSVA	•
864 2214 A 7214 845 1619 SDKGGKKADRKNHLRHAFPLLI	
DPKVPVDADHVQGQDPGRAAH	
KVSKDPLAPDEVGDTDEGHDRI	
HGHDQEEVAYEERACEGGKFA'	
DEALREAMPKVAKYAGGTNDK	
PISFAVFPNEDGSLQKKLKVWF	
PAPSDKSVKIEEREGITVYSMQF YVAQATRLRAALEGTATYRGD	
MKPYGRRNEIWLLKT	TITCIGIDIF
865 2215 A 7246 559 682 RRLGAVAHAYTSSTLGGRGGW	TT*GOELOTS
LANMAKPRLY	OQDDQ10
866 2216 A 7257 641 1310 TCTYKYLMGWIRGRRSRHSWE	MSEFHNYNL
DLKKSDFSTRWQKQRCPVVKSF	
FCCFIAVAMGIRFIIMVAIWSAV	•
QIPLTESYCGPCPKNWICYKNNO	
WYESQASCMSQNASLLKVYSKI KSYHWMGLVHIPTNGSWQWEI	
IIEMQKGDCALYASSFKGYIENC	
QRTV	
867 2217 A 7288 151 396 SIKILEAFGSNGPDFWFFRYWSP	*LFRQQVVFI
MPFFQTLWLMNANRFCSIFTTTT	NVANNCWW
TPYHCWLSVVVCRCESHGI	
868 2218 A 7298 3 272 PDTVIGGRGSGGKEFGRWVLW	
KGSCPAGGSRMVSESD*EGRGC AGS*WR*GSRPAGRGTPPRSLSI	
869 2219 A 7332 1223 332 PRRDAEDRDESCLNPAFPIGLLH	
FLTLCTWLLLGPGLLATVRAE	
YRLVRPADINFLACVMECEGKL	
KELLQLSKPELPQDGTSTLRENS	
KRYGGFMKRYGGFMKKMDELY	
NGSEILAKRYGGFMKKDAEEDI	
KELLETGDNRERSHHQDGSDNE FMRGLKRSPQLKEKAKELQKRY	
POKW*MTSPONRYGGFLKRFAE	
SYSKEVPEMEKRYGGFMRF	
870 2220 A 7382 216 1018 EIHQRLTERTQFLDESRKNPNS*(QANLLRGGG
AGQGRGREGAESGGSRGEGPGS	
FWSPRSQRRGCCGRRAPRPEAM	
TEEDPGPARGPRSGLAAYFFMG	
KGLQLLLSLLAFICEEVVSQCTL	
VSCSAFLLSLLILIVYCTPFYERV FYITLGTGCVFLLASIIFVSTHDR	
GFIASFMFLLDFITMLYEKRQES	
RAEALTEPLNA	45.00 St. 11
871 2221 A 7403 3 393 SCAMCSGLL*LLLPIWLSWTLGT	TRGSEPRSVN
DPGNMSFVKETVDKLLTGFRCF	
ALRGAALPGESEAGDPESLRSSV	MADWIQYS

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod .	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl- eotide	peptide seq-		in USSN	nucleotide location	location corresponding	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ł	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
ŀ	ŀ			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		1		peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
			:			DLWEAEVSTPRCEAGFCQECFRTPGNQEKDG
872	2222	A	7413	1061	250	PPIC
8/2	2222	A	/413	1001	359	FVDIVSVVEFPHCPEARFPAQHGQDSKRLTLC PGGS*PQATLHLDRMRVSASPTKEIQVKKYK
			! !	! !		CGLIKPCPANYFAFKICSGAANVVGPTMCFED
Ì			ł	Ì	1	RMIMSPVKNNVGRGLNIALVNGTTGAVLGQ
						KAFDMYSGDVMHLVKFLKEIPGGALVLVAS
			!			YDDPGTKMNDESRKLFSDLGSSYAKQLGFRD
j]	}]	j	j	SWVFIGAKDLRGKSPFEQFLKEQPQTQNKYE
						GWPELLEMEGCMPPKPF
873	2223	Α	7429	2242	2394	ILKCAGHGGSCL*SQHFGRLRWEDRLRLGVQ
074	0004		- 150	:		DHPGQHCETPSLLKIERKLF
874	2224	Α	7468	146	894	PCTSCVLWATLHLPASTRKAPQAECGMISITE
			!	-		WQKIGVGITGFGIFFILFGTLLYFDSVLLAFGN
			i '			LLFLTGLSLIIGLRKTFWFFFQRHKLKGTSFLL GGVVIVLLRWPLLGMFLETYGFFSLFKGFFPV
		ĺ	1	l		AFGFLGNVCNIPFLGALFRRLQGTSSMV*KTE
		i				MSSLNLDHWLKGAKREEWEPPPQSPALTHSP
		ļ				TYPGPPQVQKERNGAEQLTSNPQVDSRGCQE
						AEMQTPRRLGWGWYHTLTLYLWEEK
875	2225	A	7498	91	251	GEKPVPTWLQDEAGQWLLGFVAQPWGWPG
				<u> </u>		SERHEP*HGGVLFRLGPSAPPGKL
876	2226	A	7544	403	587	YSCLCFLFKHITSFKNSVHIWLGTVVHAYNPN
877	2227		7566		0.10	ILGGQGGWIA*GQEFKTSLGNTVRPCLYK
0//	2221	A	7566	2	940	GCAPDTRFFVPEPGGRGAAPWVALVARGGC TFKDKVLVAARRNASAVVLYNEERYGNITLP
i i		ľ	ľ			MSHAGTGNIVVIMISYPKGREILELVQKGIPV
					:	TMTIGVGTRHVQEFISGQSVVFVAIAFITMMII
			i			SLAWLIFYYIQRFLYTGSQIGSQSHRKETKKVI
					l	GQLLLHTVKHGEKGIDVDAENCAVCIENFKV
						KDIIRIL PCKHIFHRICIDPWLLDHRTCPMCKL
		ŀ	,			DVIKALGYWGEPGDVQEMPAPESPPGRDPAA
						NLSLALPDDDGSDESSPPSASPAESEPQCDPSF
878	2228	A	7586	315	1232	KGDAGENTALLEAGRSDSRHGGPIS ERSLLCKVDVRWIYVSEGTKTQRRHRQGSLR
0/0	2220	A	/380	313	1232	RGRMQAACWYVLFLLQPTVYLVTCANLTNG
						GKSELLKSGSSKSTLKHIWTESSKDLSISRLLS
1						QTFRGKENDTDLDLRYDTPEPYSEQDLWDW
1						LRNSTDLQEPRPRAKRRPIVKTGKFKKMFGW
		•				GDFHSNIKTVKLNLLITGKIVDHGNGTFSVYF
						RHNSTGQGNVSVSLVPPTKIVEFDLAQQTVID
] ,						AKDSKSFNCRIEYEKVDKATKNTLCNYDPSK
					·	TCYQEQTQSHVSWLCSKPFKVICIYISFYSTD
879	2229	Α	7605	170	201	YKLVQKVCPDYNYHSDTPYFPSG
880	2229	A	7605 7612	93	391	TESWKLKWWSPTCLDQLNGSAPGNVFIHG
000	2230	A	/012	73	659	DAAVAMTAQGGLVANRGRRFKWAIELSGPG GGSRGRSDRGSGQGDSLYPVGYLDKQVPDTS
	.					VQETDRILVEKRCWDIALGPLKQIPMNLFIMY
			}	ı		MAGNTISIFPTMMVCMMAWRPIQALMAISAT
					•	FKMLESSSQKFLQGLVYLIGNLMGLALAVYK
						CQSMGLLPTHASDWLAFIEPPERMEFSGGGL
						LL
881	2231	A	7615	291	1452	SPQKTMRSHTITMTTTSVSSWPYSSHRMRFIT
						NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT
						SYSVIFIVGLVGNIIALYVFLGIHRKRNSIQIYL
	İ			l	' I	LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL
				Į.		CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI
				{	1	QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL TLKKGGHNSTMCFHYRDKHNAKGEAIFNFIL
L						TEATTOOTH OF INCIDENT ADAMNANUE MICHEL

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion VVMFWLIFLLIILSYIKIGKNLLRISKRRSKFPN
						SGKYATTARNSFIVLIIFTICFVPYHAFRFIYISS QLNVSSCYWKEIVHKTNEIMLVLSSFNSCLDP VMYFLMSSNIRKIMCQLLFRRFQGEPSRSEST SEFKPGYSLHDTSVAVKIQSSSKST
882	2232	A	7617	67	379	RQMALLKANKDLISAGLKEFSVLLNQQVFND PLVSEEDMVTVVEDWMNFYINYYRQQVTGE PQERDKALQELRQELNTLANPFLAKYRDFLK SHELPSHPPPSS
883	2233	A	7622	400	215	KVKTCRYNPKYSAANDTGFVDIPSREKDLAK AVATVGPISVAVGASHVFFQFYKKGKHLSS
884	2234	A	7638	2640	2861	APVLILQMVKLSIVLTPQFLSHDQGQLTKELQ QHVKSVTCPCEYLRKVSECRQMGPGALEQFP GLSCHTSHSG
885	2235	A	7642	201	455	PSRGKMELEAMSRYTSPVNPAVFPHLTVVLL AIGMFFTAWFFVYEVTSTKYTRDIYKELLISL VASLFMGFGVLFLLLWVGIYV
886	2236	A	7692	61	569	APENPFSRQHFNSETKVKLSLKTGTWLGNHA HLGEHFSTHHELGLSGKVVGFLVKNILEVIRN GGMETRHPGKVSSWFHRWDSRAEQHNHAE HHEDVPQGDEDSKVSEAQQEFPDVVTCAGLP GLLPKALRVLLFQLKVQHRPGIHQQRPEQQD VSDHRYGRSVRQNRK
887	2237	A	7693	85	315	NPGCCLPVAMRTSYLLLFTLCLLLSEMASGG NFLTGLGHRSDHYNCVSSGGQCLYSACPIFTK IQGTCYRGKAKCCK
888	2238	A	7702	242	1298	APSHRRRYLSPSRSAGQLGNMALERLCSVLK VLLITVLVVEGIAVAQKTQDGQNIGIKHIPAT QCGIWVRTSNGGHFASPNYPDSYPPNKECIYI LEAAPRQRIELTFDEHYYIEPSFECRFDHLEVR DGPFGFSPLIDRYCGVKSPPLIRSTGRFMWIKF SSDEELEGLGFRAKYSFIPDPDFTYLGGILNPIP DCQFELSGADGIVRSSQVEQEEKTKPGQAVD CIWTIKATPKAKIYLRFLDYQMEHSNECKRNF VAVYDGSSSIENLKAKFCSTVANDVMLKTGI GVIRMWADEGSRLNRFRMLFTSFGGASPAQA ALSFCHSNMCINNSLVCNGVQNCAYPWDEN HC
889	2239	A	7707	185	2911	CHYIMNPSTHIPASAGGSILGLFDFFGLGLGE MTMDALLARLKLLNPDDLREEIVKAGLKCGP ITSTTRFIFEKKLAQALLEQGGRLSSFYHHEA GVTALSQDPQRILKPAEGNPTDQAGFSEDRDF GYSVGLNPPEEEAVTSKTCSVPPSDTDTYRAG ATASKEPPLYYGVCPVYEDVPARNERIYVYE NKKEALQAVKMIKGSRFKAFSTREDAEKFAR GICDYFPSPSKTSLPLSPVKTAPLFSNDRLKDG LCLSESETVNKERANSYKNPRTQDLTAKLRK AVEKGEEDTFSDLIWSNPRYLIGSGDNPTIVQ EGCRYNVMHVAAKENQASICQLTLDVLENP DFMRLMYPDDDEAMLQKRIRYVVDLYLNTP DKMGYDTPLHFACKFGNADVVNVLSSHHLI VKNSRNKYDKTPEDVICERSKNKSVELKERIR EYLKGHYYVPLLRAEETSSPVIGELWSPDQTA EASHVSRYGGSPRDPVLTLRAFAGPLSPAKAE DFRKLWKTPPREKAGFLHHVKKSDPERGFER VGRELAHELGYPWVEYWEFLGCFVDLSSQE GLQRLEEYLTQQEIGKKAQQETGEREASCRD KATTSGSNSISVRAFLDEDDMSLEEIKNRQNA ARNNSPPTVGAFGHTRCSAFPLEQEADLIEAA

890 2240 A 7711 360 269 RHMPVIPALWEAEVGGLLEPRSSRSAWA 891 2241 A 7721 61 1175 KLPWEPSFLIKMQIIRHSEQTLKTALISKN VSQYEKLDAGEQRLMNEAFQPASDLFGI HSPSDWITSHPEAPQDFEQFFSDPYRKTP KRSIYIQSIGSLGNTRIISEEVIKWLTGYCI YGLRVKLLEPVPVSVTRCSFRVNENTHN AGDILKFLKKKKPEDAFCVVGITMIDLYI WNFVFGQASLTDGVGIFSFARYGSDFYS KGKVKKLKKTSSSDYSIFDNYYIPEITSV SCKTLTHEIGHIFGLRHCQWLACLMQGS EEADRRPLNLCPICLHKLQCAVGFSIVER LVRWIDDESSDTPGATPEHSHEDNGNLP EAFKEWKEWIIKCLAVLQK 892 2242 A 7723 2 1650 SAPTAPARPCRAERGSGGGMLALLAASV VAAGAQDSPAPGSRFVCTALPPEAVHAC PAMPMQGGAQSPEELRAAVLQLRETV KETLASARAIRELTGKLARCEGLAGGKA GATGKDTMGDLPRDPGHVVEQLSRSLQ DRLESLEPLPAMPMQGGAQSPEEELRAA LRETVVQQKETLASARAIRELTGKLARC AGGKARGAGATGKDTMGDLPRDPGHV	APR KTP OLG OLEV ADV PAKP
VSQYEKLDAGEQRLMNEAFQPASDLFGI HSPSDWITSHPEAPQDFEQFFSDPYRKTP KRSIYIQSIGSLGNTRIISEEYIKWLTGYCI YGLRVKLLEPVPVSVTRCSFRVNENTHN AGDILKFLKKKKPEDAFCVVGITMIDLYI WNFVFGQASLTDGVGIFSFARYGSDFYS KGKVKKLKKTSSSDYSIFDNYYIPEITSV SCKTLTHEIGHIFGLRHCQWLACLMQGS EEADRRPLNLCPICLHKLQCAVGFSIVER LVRWIDDESSDTPGATPEHSHEDNGNLP EAFKEWKEWIIKCLAVLQK 892 2242 A 7723 2 1650 SAPTAPARPCRAERGSGGMLALLAASV VAAGAQDSPAPGSRFVCTALPPEAVHAC PAMPMQGGAQSPEEELRAAVLQLRETV KETLASARAIRELTGKLARCEGLAGGKA GATGKDTMGDLPRDPGHVVEQLSRSLQ DRLESLEPLPAMPMQGGAQSPEEELRAA LRETVVQQKETLASARAIRELTGKLARC	TE
892 2242 A 7723 2 1650 SAPTAPARPCRAERGSGGMLALLAASV VAAGAQDSPAPGSRFVCTALPPEAVHAC PAMPMQGGAQSPEEELRAAVLQLRETV KETLASARAIRELTGKLARCEGLAGGKA GATGKDTMGDLPRDPGHVVEQLSRSLQ DRLESLEPLPAMPMQGGAQSPEEELRAA LRETVVQQKETLASARAIRELTGKLARC	ITL SPN AYF LQIH RDS MHY LLR VHL
LSRSLQTLKDRLESLEHQLRANVSNAGL FREVLQORLGELERQLLRKGAELEDEKS	CPL /QQ RGA LK VLQ GGL /EQ PGD
NETSAHRQKTESTLNALLQRVTELERGN KSPNAFKVSLPLRTNYLYGKIKKTLPELY ICLWLRSSASPGMGTPFSYAVPGQANEIV WGNNPIELLINDKVAQLPLFVSDGKWHI TWTTRDGMWEAFQDGKKLGTGENLAPY KPGGVLILGQEQDTVGGRFDATQAFVGE FNIWDRVLRAQEIVNIANCSTNMPGNIIPY NNVDVFGGASKWPVETCEERLLDL 893 2243 A 7729 3554 2419 LTAGTAMNYPLTI FMDLENI FDLEWELL	AFT LIE ICV VHPI LSQ VVD
893 2243 A 7729 3554 2419 LTAGTAMNYPLTLEMDLENLEDLFWELT DNYNDTSLVENHLCPATEGPLMASFKAN VAYSLIFLLGVIGNVLVLVILERHRQTRS: FLFHLAVADLLLVFILPFAVAEGSVGWV. LCKTVIALHKVNFYCSSLLLACIAVDRYI HAVHAYRHRRLLSIHITCGTIWLVGFLLA LFAKVSQGHHNNSLPRCTFSQENQAETH TSRFLYHVAGFLLPMLVMGWCYVGVVF QAQRRPQRQKAVRVAILVTSIFFLCWSPY IFLDTLARLKAVDNTCKLNGSLPVAITMU GLAHCCLNPMLYTFAGVKFRSDLSRLLT CTGPASLCQLFPSWRRSSLSESENATSLT	FVP TET GTF AIV LPEI AWF RLR HIV EFL KLG
894 2244 A 7738 670 287 FVTRAGRWGAGARVRGGAGGMASGAA VLAPVRSGALRSGPSLRKDGDVSAAWSC SLVPSRSVIVTRSGAILPKPVKMSFGLLRV VIPFLYVGTLISKNFAALLEEHDIFVPEDD D	RWL SGR FSI
895 2245 A 7753 119 278 APYAHSQVHCLDKVCGLLPFLNPEVPDQ	YR
LWLSLFLHAGKEAPHCPRTRPL 896 2246 A 7754 1 372 SPAWWNSQQRVVSPFLALLTLEPTFHHILI	PIM

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion QVSTAALAVLLCTMALCNQVLSAPLAADTPT ACCFSYTSRQIPQNFIADYFETSSQCSKPSVIFL
897	2247	A	7761	1725	445	TKRGRQVCADPSEEWVQKYVSDLELSA RPRRGTHHFSCVLGSFRVSAMFPRVSTFLPL RPLSRHPLSSGSPETSAAAIMLLTVRHGTVRY RSSALLARTKNNIQRYFGTNSVICSKKDKQSV RTEETSKETSESQDSEKENTKKDLLGIIKGMK VELSTVNVRTTKPPKRRPLKSLEATLGRLRRA TEYAPKKRIEPLSPELVAAASAVADSLPFDKQ TTKSELLSQLQQHEEESRAQRDAKRPKISFSNI ISDMKVARSATARVRSRPELRIQFDEGYDNYP GQEKTDDLKKRKNIFTGKRLNIFDMMAVTKE APETDTSPSLWDVEFAKQLATVNEQPLQNGF EELIQWTKEGKLWEFPINNEAGFDDDGSEFH EHIFLEKHLESFPKQGPIRHFMELVTCGLSKNP YLSVKQKVEHIEWFRNYFNEKKDILKESNIQF KLRPWKFLFRNN
898	2248	A	7775 '	85	496	SCOTTOPPAQSCSTGTMRIMLLFTAILAFSLA QSFGAVCKEPQEEVVPGGGRSKRDPDLYQLL QRLFKSHSSLEGLLKALSQASTDPKESTSPEK RDMHDFFVGLMGKRSVQPDSPTDVNQENVP SFGILKYPPRAE
899	2249	A	7785	179	703	PFHLGASSNTFRLQVQTQESKAQKEVKMGFI FSKSMNESMKNQKEFMLMNARLQLERQLIM QSEMRERQMAMQIAWSREFLKYFGTFFGLA AISLTAGAIKKKKPAFLVPIVPLSFILTYQYDL GYGTLLERMKGEAEDILETEKSKLQLPRGMIT FESIEKARKEQSRFFIDK
900	2250	A	7789	1465	300	VWLPLKSYKIRSPSLHCQCEIFREEFLFSSLQE GRDKDTFSKMAMVSEFLKQAWFIENEEQEY VQTVKSSKGGPGSAVSPYPTFNPSSDVAALH KAIMVKGVDEATIIDILTKRNNAQRQQIKAAY LQETGKPLDETLKKALTGHLEEVVLALLKTP AQFDADELRAAMKGLGTDEDTLIEILASRTN KEIRDINRVYREELKRDLAKDITSDTSGDFRN ALLSLAKGDRSEDFGVNEDLADSDARALYEA GERRKGTDVNVFNTILTTRSYPQLRVFQKY TKYSKHDMNKVLDLELKGDIEKCLTAIVKCA TSKPAFFAEKLHQAMKGVGTRHKALIRIMVS RSEIDMNDIKAFYQKMYGISLCQAILDETKGD YEKILVALCGGN
901	2251	A	7796	2	807	VEFHPQRARAGARAPSMGVLLTQRTLLSLVL ALLFPSMASMAAIGSCSKEYRVLLGQLQKQT DLMQDTSRLLDPYIRIQGLDVPKLREHCRERP GAFPSEETLRGLGRRCFLQTLNATLGCVLHRL ADLEQRLPKAQDLERSGLNIEDLEKLQMARP NILGLRNNIYCMAQLLDNSDTAEPTKAGRGA SQPPTPTPASDAFQRKLEGCRFLHGYHRFMH SVGRVFSKWGESPNRSRRHSPHQALRKGVRR TRPSRKGKRLMTRGQLPR
902	2252	A	7802	2	721	TAARRRQKGTAARRLQKGTAARRRQKGTAA RRRQKGTAARRPQKGTAARRRQKGTAARRR QKGTAARRRQKGTAARRPQKGTAARRRQKG TAARRRQKGTAARRRQKGLAIASRGCPCASR AGGVRGAGSRLRAMAPKVFRQYWDIPDGTD CHRKAYSTTSIASVAGLTAAAYRVTLNPPGTF LEGVAKVGQYTFTAAAVGAVFGLTTCISAHV REKPDDPLNYFLGGCAGGLTLGARTHNYGIG AAACVYFGIAASLVKMGRLEGWEVFAKPKV

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ĺ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first amino acid	acid residue of peptide	Q=Glutamine, R=Arginine, S=Serine,
		i		residue of	sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
		[peptide	sequence	/=possible nucleotide deletion, \=possible
1		}	l	sequence		nucleotide insertion
903	2253	Α	7807	1	584	PWLPWSDGRAARSSRKCPRSRFPVQVGKMA
						VSTVFSTSSLMLALSRHSLLSPLLSVTSFRRFY
		1	1			RGDSPTDSQKDMIEIPLPPWQERTDESIETKR
						ARLLYESRKRGMLENCILLSLFAKEHLQHMT
			i			EKQLNLYDRLINEPSNDWDIYYWATEAKPAP
						EIFENEVMALLRDFAKNKNKEQRLRAPDLEY
004	0054		7017	10		LFEKPR
904	2254	A	7813	40	821	GAGRALGHLETGAGDVAAALPARKFPRSLLG
						AGARLTGWTMNVFRILGDLSHLLAMILLLGK
						IWRSKCCKGISGKSQILFALVFTTRYLDLFTNF
1		ł				ISIYNTVMKVVFLLCAYVTVYMIYGKFRKTF DSENDTFRLEFLLVPVIGLSFLENYSFTLLEIL
•						WTFSIYLESVAILPOLFMISKTGEAETITTHYL
						FFLGLYRALYLANWIRRYQTENFYDQIAVVS
1.			1			GVVQTIFYCDFFYLYVTKGRSWDDSNADTGL
						RSYSSI
905	2255	A	7817	1399	881	LSNKDVLSPQLKDENSKLRRKLNEVQSFSEA
						QTEMVRTLERKLEAKMIKEESDYHDLESVVQ
]			QVEQNLELMTKRAVKAENHVVKLKQEISLL
						QAQVSNFQRENEALRCGQGASLTVVKQNAD
1						VALQNLRVVMNSAQASIEQLVSGAETLNLVA
906	2256	A	7822		1460	EILKSIDRISEVKDEEEDS
300	2230	A	/822	3	1462	DSPRNRFEILGRPTRTPTRPGPRPAMEDLDAL
1						LSDLETTTSHMPRSGAPKERPAEPLTPPPSYG HQPQTGSGESSGASGDKDHLYSTVCKPRSPK
1						PAAPAAPPFSSSSGVLGTGLCELDRLLQELNA
						TQFNITDEIMSQFPSSKVASGEQKEDQSEDKK
1 1						RPSLPSSPSPGLPKASATSATLELDRLMASLSD
						FRVQNHLPASGPTQPPVVSSTNEGSPSPPEPTG
						KGSLDTMLGLLQSDLSRRGVPTQAKGLCGSC
1						NKPIAGQVVTALGRAWHPEHFVCGGCSTAL
						GGSSFFEKDGAPFCPECYFERFSPRCGFCNQPI
						RHKMVTALGTHWHPEHFCCVSCGEPFGDEG
1 1						FHEREGRPYCRRDFLQLFAPRCQGCQGPILDN
						YISALSALWHPDCFVCRECFAPFSGGSFFEHE GRPLCENHFHARRGSLCATCGLPVTGRCVSA
						LGRRFHPDHFTCTFCLRPLTKGSFQERAGKPY
1						COPCFLKLFG
907	2257	A	7828	1792	1671	FIYVNQSFAPSPDQEVGTLYECFGSDGKLVLH
						YCKSQAWG
908	2258	A	7842	110	1172	KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG
					ŀ	SLQPPPSGLKQSSHLSLSSSWDFRHAPTHPET
				ſ	1	YTCPKMIEMEQAEAQLAELDLLASMFPGENE
						LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI
						NMNLDVSDEKMAMFSLACILPFKYPAVLPEI
					. '	TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV
1	1	l		1	;	CILNATEWVRÉHASGYVSRDTSSSPTTGSTVQ SVDLIFTRLWIYSHHIYNKCKRKNILEWAKEL
						SLSGFSMPGKPGVVCVEGPQSACEEFWARLR
						KLNWKRILIRHREDIPFDGTNDETERQRKFSIF
		ļ	ļ	ļ	.,	EEKVFSVNGARGNHMDFGOLYOFLNTKGCG
		}	i			DVFQMFLWV
909	2259	A	7870	3067	2923	EGICVYTFIYVHMYTRTCMHTYPYMYMNSV
						LISSEILLIPSKYLFESK
910	2260	A	7884	212	4874	GALTWSHPLLAVCPQGVWLGSTPSGSPALLP
	i	ĺ	Í	- 1		PSHRVNAEPGCVVTNACASGPCPPHANCRDL
			ļ		1	WQTFSCTCQPGYYGPGCVDACLLNPCQNQG
		!				SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD
L	i	i				QQCPRGWWGSPTCGPCNCDVHKGFDPNCNK

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl- eotide	peptide seq-		in USSN	nucleotide location	location corresponding	F=Phenylalanine, G=Glycine, H=Histidine, l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1		ţ		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		,		peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
						TNGQCHCKEFHYRPRGSDSCLPCDCYPVGST
						SRSCAPHSGQCPCRPGALGRQCNSCDSPFAEV
						TASGCRVLYDACPKSLRSGVWWPQTKFGVL
						ATVPCPRGALGLRGAGAAVRLCDEAQGWLE PDLFNCTSPAFRELSLLLDGLELNKTALDTME
						AKKLAQRLREVTGHTDHYFSQDVRVTARLL
1					•	AHLLAFESHQQGFGLTATQDAHFNENLLWA
						GSALLAPETGDLWAALGQRAPGGSPGSAGLV
						RHLEEYAATLARNMELTYLNPMGLVTPNIML
						SIDRMEHPSSPRGARRYPRYHSNLFRGQDAW
						DPHTHVLLPSQSPRPSPSEVLPTSSSIENSTTSS
						VVPPPAPPEPEPGISIILLVYRTLGGLLPAQFQ
1 1						AERRGARLPONPVMNSPVVSVAVFHGRNFLR
						GILESPISLEFRLLQTANRSKAICVQWDPPGLA EQHGVWTARDCELVHRNGSHARCRCSRTGT
						FGVLMDASPRERLEGDLELLAVFTHVVVAVS
						VAALVLTAAILLSLRSLKSNVRGIHANVAAA
						LGVAELLFLLGIHRTHNQLVCTAVVILLHYFF
					,	LSTFAWLFVQGLHLYRMQVEPRNVDRGAMR
						FYHALGWGVPAVLLGLAVGLDPEGYGNPDF
			' I		' I	CWISVHEPLIWSFAGPVVLVIVMNGTMFLLA
						ARTSCSTGQREAKKTSALTLRSSFLLLLLVSA
]					i	SWLFGLLAVNHSILAFHYLHAGLCGLQGLAV LLLFCVLNADARAAWMPACLGRKAAPEEAR
]]						PAPGLGPGAYNNTALFEESGLIRITLGASTVSS
1	Ì				' i	VSSARSGRTQDQDSQRGRSYLRDNVLVRHGS
					*	AADHTDHSLQAHAGPTDLDVAMFHRDAGA
						DSDSDSDLSLEEERSLSIPSSESEDNGRTRGRF
	٠]		QRPLCRAAQSERLLTHPKDVDGNDLLSYWPA
				ľ	.	LGECEAAPCALQTWGSERRLGLDTSKDAAN
		İ				NNQPDPALTSGDETSLGRAQRQRKGILKNRL
				İ		QYPLVPQTRGAPELSWCRAATLGHRAVPAAS YGRIYAGGGTGSLSQPASRYSSREQLDLLLRR
1	l	1	1		1	QLSRERLEEAPAPVLRPLSRPGSQECMDAAPG
						RLEPKDRGSTLPRRQPPRDYPGAMAGRFGSR
					,	DALDLGAPREWLSTLPPPRRTRDLDPQPPPLP
						LSPQRQLSRDPLLPSRPLDSLSRSSNSREQLDQ
1		- 1	ļ			VPSRHPSREALGPLPQLLRAREDSVSGPSHGP
			ļ		j	STEQLDILSSILASFNSSALSSVQSSSTPLGPHT
						TATPSATASVLGPSTPRSATSHSISELSPDSEPR DTQALLSATQAMDLRRRDYHMERPLLNOEH
[1				LEELGRWGSAPRTHQWRTWLQCSRARAYAL
	ł		ŀ		, }	LLQHLPVLVWLPRYPVRDWLLGDLLSGLSVA
	i	1			•	IMQLPQGLAYALLAGLPPVFGLYSSFYPVFIY
						FLFGTSRHISVESLCVPGPVDT
911	2261	A	7890	21	806	EFGTSRSSRSMAEDLGLSFGETASVEMLPEHG
		[SCRPKARSSSARWALTCCLVLLPFLAGLTTYL
		١		j		LVSQLRAQGEACVQFQALKGQEFAPSHQQV
		l				YAPLRADGDKPRAHLTVVRQTPTQHFKNQFP
	ł	- 1	1	1	1	ALHWEHELGLAFTKNRMNYTNKFLLIPESGD YFIYSQVTFRGMTSECSEIROAGRPNKPDSITV
					1	VITKVTDSYPEPTQLLMGTKSVCEVGSNWFQ
	1	J				PIYLGAMFSLOEGDKLMVNVSDISLVDYTKE
	İ	ļ	-	i	ļ	DKTFFGAFLL
912	2262	A	7891	1263	111	ACGIRHEGALPGLTATPEAMLRFLPDLAFSFL
			ļ	ĺ		LILALGQAVQFQEYVFLQFLGLDKAPSPQKFQ
]	l				PVPYILKKIFQDREAAATTGVSRDLCYVKELG
,	-	ł		1	1	VRGNVLRFLPDQGFFLYPKKISQASSCLQKLL
	}		J	1	į	YFNLSAIKEREQLTLAQLGLDLGPNSYYNLGP
						ELELALFLVQEPHVWGQTTPKPGKMFVLRSV

	SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
ĺ					peptide sequence		/=possible nucleotide deletion, \=possible nucleotide insertion
							PWPQGAVHFNLLDVAKDWNDNPRKNFGLFL EILVKEDRDSGVNFQPEDTCARLRCSLHASLL
		!					VVTLNPDQCHPSRKRRAAIPVPKLSCKNLCH RHQLFINFRDLGWHKWIIAPKGFMANYCHGE
							CPFSLTISLNSSNYAFMQALMHAVDPEIPQAV CIPTKLSPISMLYQDNNDNVILRHYEDMVVD
ŀ	913	2263	Α	7892	15	849	ECGCG ASRLPRGPGCGADMRPLLGLLLVFAGCTFAL
				,0,2		0.5	YLLSTRLPRGRRLGSTEEAGGRSLWFPSDLAE
					,	ĺ	LRELSEVLREYRKEHQAYVFLLFCGAYLYKQ
							GFAIPGSSFLNVLAGALFGPWLGLLLCCVLTS VGATCCYLLSSIFGKQLVVSYFPDKVALLQR
1							KVEENRNSLFFFLLFLRLFPMTPNWFLNLSAPI
							LNIPIVQFFFSVLIGLIPYNFICVQTGSILSTLTS
							LDALFSWDTVFKLLAIAMVALIPGTLIKKFSQ KHLQLNETSTANHIHSRKDT
t	914	2264	Α	7893	815	959	KSGWVWWLTPLIPALWEAQTEGSLRPEVKN
-	915	2265	_	7000			RLSNITRPFFSKKKKILV
ĺ	913	2265	A	7909	3	641	HASGPGGLLRRRRGSGANMPVARSWVCRKT YVTPRRPFEKSRLDOELKLIGEYGLRNKREV
۱							WRVKFTLAKIRKAARELLTLDEKDPRRLFEG
1						ł	NALLRRLVRIGVLDEGKMKLDYILGLKIEDFL
1							ERRLQTQVFKLGLAKSIHHAHVLIQQCHIRVR EQVVNILFFTVRLDSQKHIDFSLCFPIGVANPS
							HVKRKNASKGQGGAGARDDEEEE
r	916	2266	Α	7914	3	967	VAHTQWHTCQRLSQLTHRSILKYLLIDTHAC
							QVLILKHTHASLSLPSCQECFPSSIPSASHMVS
							HPHPPPSPRWGQTPEGLPAASPCGPGPRSCFS SILPTGDSWGMLACLCTVLWHLPAVPALNRT
				-			GDPGPGPSIQKTYDLTRYLEHQLRSLAGTYLN
4							YLGPPFNEPDFNPPRLGAETLPRATVDLEVW
	l						RSLNDKLRLTQNYEAYSHLLCYLRGLNRQAA TAELRRSLAHFCTSLQGLLGSIAGVMAALGY
		ŀ					PLPQPLPGTEPTWTPGPAHSDFLQKMDDFWL
							LKELQTWLWRSAKDFNRLKKKMQPPAAAVT
+	917	2267	A	7921	2	1166	LHLGAHGF RRPRGOOL VOEVOTENIVEVA POOVA FITTORI
ı)., 	2207	^	/921	2	1100	RPRRGQGLVQEVQTENVTVAEGGVAEITCRL HQYDGSIVVIQNPAROTLFFNGTRALKDERFO
							LEEFSPRRVRIRLSDARLEDEGGYFCQLYTED
١							THHQIATLTVLVAPENPVVEVREQAVEGGEV
	Ī		1				ELSCLVPRSRPAATLRWYRDRKELKGVSSSQ ENGKVWSVASTVRFRVDRKDDGGIIICEAQN
1							QALPSGHSKQTQYVLDVQYSPTARIHASQAV
Į	į		- 1				VREGDTLVLTCAVTGNPRPNQIRWNRGNESL
İ							PERAEAVGETLTLPGLVSADNGTYTCEASNK HGHARALYVLVVYGESRLRPTEGGGGAPDP
							GAVVEAQTSVPYAIVGGILALLVFLIICVLVG
1	1						MVWCSVRQKGSYLTHEASGLDEQGEAREAF
-	918	2268	$\overline{\mathbf{A}}$	7938	3	2653	LNGSDGHKRKEEFFI
	-10	-200	^	1230	,	2033	RRRLPPASPPSSSVSSSLSPSAVVMACRWSTK ESPRWRSALLLLFLAGVYGNGALAEHSENVH
							ISGVSTACGETPEQIRAPSGIITSPGWPSEYPAK
						ĺ	INCSWFIRANPGEIITISFQDFDIQGSRRCNLD
	1						WLTIETYKNIESYRACGSTIPPPYISSQDHIWIR FHSDDNISRKGFRLAYFSGKSEEPNCACDQFR
	1	- 1					CGNGKCIPEAWKCNNMDECGDRSDEEICAKE
							ANPPTAAAFQPCAYNQFQCLSRFTKVYTCLP
	1		ļ				ESLKCDGNIDCLDLGDEIDCDVPTCGQWLKY
L		1					FYGTFNSPNYPDFYPPGSNCTWLIDTGDHRK

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /-possible nucleotide deletion, \=possible nucleotide insertion VILRFTDFKLDGTGYGDYVKIYDGLEENPHK LLRVLTAFDSHAPLTVVSSSGQIRVHFCADKV NAARGFNATYQVDGFCLPWEIPCGGNWGCY TEQQRCDGYWHCPNGRDETNCTMCQKEEFP CSRNGVCYPRSDRCNYQNHCPNGSDEKNCFF CQPGNFHCKNNRCVFESWVCDSQDDCGDGS DEENCPVIVPTRVITAAVIGSLICGLLLVIALG CTCKLYSLRMFERRSFETQLSRVEAELLRREA PPSYGQLIAQGLIPPVEDFPVCSPNQASVLENL
919	2269	A	7951	1674	1839	RLAVRSQLGFTSVRLPMAGRSSNIWNRIFNFA RSRHSGSLALVSADGDEVVPSQSTSREPERNH THRSLFSVESDDTDTENERRDMAGASGVAA PLPQKVPPTTAVEATVGACASSSTQSTRGGH ADNGRDVTSVEPPSVSPARHQLTSALSRMTQ GLRWVRFTLGRSSSLSQNQSPLRQLDNGVSG REDDDDVEMLIPISDGSSDFDVNDCSRPLLDL ASDQGQGLRQPYNATNPGVRPSNRDGPCERC GIVHTAQIPDTCLEVTLKNETSDDEALLLC VVRVTCCPPARSTTERTNAYDEEDCVEMVAS
920	2270	A	7953	47	572	GGWNDVACHTTMYFMCEFDKKNM GGRASWPEQAKEPRREGHTDKQQTEDVLAA GLRCLPHLPAICARRMSPAFRAMDVEPRAKG VLLEPFVHQVGGHSCVLRFNETTLCKPLVPRE HQFYETLPAEMRKFTPQYKGKSQLLEGLPHW RGDVRDRGHGRPWQPSLEPSLPPTLCFPSLSS FSSSWPSAQHLTPSVFNPW
921	2271	A	7957	612	812	RSGRTVVTGIGYSKALQSSNRNTKSLLQNEF MMVYSFRALSFKESTWATFQHGGEATKSRSL SSTQ
922	2272	A	7967	1443	1660	ENITEKWKEIWMCRGNKKSCCWTFIKDRHLT VSCCKSKSGETLLICIFCSNLVGFFFFGIRGFSN WELVKPN
923		A	7981		3023	GSAPRAATAMARARPPPPPSPPPGLLPLLPLL LLPLLLLPAGCRALEETLMDTKWVTSELAWT SHPESGWEEVSGYDEAMNPIRTYQVCNVRES SQNNWLRTGFIWRRDVQRVYVELKFTVRDC NSIPNIPGSCKETFNLFYYEADSDVASASSPFW MENPYVKVDTIAPDESFSRLDAGRVNTKVRS FGPLSKAGFYLAFQDQGACMSLISVRAFYKK CASTTAGFALPFETLTGAEPTSLVIAPGTCIPN AVEVSVPLKLYCNGDGEWMVPVGACTCATG HEPAAKESQCRPCPPGSYKAKQGEGPCLPCPP NSRTTSPAASICTCHNNFYRADSDSADSACTT VPSPPRGVISNVNETSLILEWSEPRDLGVRDD LLYNVICKKCHGAGGASACSRCDDNVEFVPR QLGLSEPRVHTSHLLAHTRYTFEVQAVNGVS GKSPLPPRYAAVNITINQAAPSEVPTLRLHSS SGSSLTLSWAPPERPNGVILDYEMKYFEKSEG IASTVTSQMNSVQLDGLRPDARYVVQVRART VAGYGQYSRPAEFETTSERGSGAQQLQEQLP LIVGSATAGLVFVVAVVVIAIVCLRKQRHGS DSEYTEKLQQYIAPGMKVYIDPFTYEDPNEA VREFAKEIDVSCVKIEEVIGAGEFGEVCRGRL KQPGRREVFVAIKTLKVGYTERQRRDFLSEA SIMGQFDHPNIIRLEGVVTKSRPVMILTEFME NCALDSFLRLNDGQFTVIQLVGMLRGIAAGM KYLSEMNYVHRDLAARNILVNSNLVCKVSDF GLSRFLEDDPSDPTYTSSLGGKIPIRWTAPEAI AYRKFTSASDVWSYGIVMWEVMSYGERPY

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion WDMSNQDVINAVEQDYRLPPPMDCPTALHQ LMLDCWVRDRNLRPKFSQIVNTLDKLIRNAA SLKVIASAQSGMSQPLLDRTVPDYTTFTTVGD WLDAIKMGRYKESFVSAGFASFDLVAQMTA EDLLRIGVTLAGHQKKILSSIQDMRLQMNQT
924	2274	A	7985	1	503	LPVQV FRPRTKKATAMYLEHYLDSIENLPCELQRNF QLMRELDQRTEDKKAEIDILAAEYISTVKTLS PDQRVERLQKIQNAYSKCKEYSDDKVQLAM QTYEMVDKHIRRLDADLARFEADLKDKMEG SDFESSGGRGLKKGRGQKEKRGSRGRGRRTS EEDTPKKKKHKGG
925	2275	A	7994	447	589	LPCSFCAQCMSSFERVWLQQSHFHNPRWNSR
926	2276	A	7996	925	582	SPIRCYCQHWPHCVHC GPCKVCCITLAIMLQCHSFYRKDVQVEHPKS LNPKYSQIENFLSADMALKRKCLLSISDLDFW IWDAQPVGIMQTLQNLKKIPNPGCFWSQAFQI RDTQPILPLGGRYYITIRQ
927	2277	A	7998	2	353	RIQRPLNSRSPNHSLFVKAELTAKQATMKLSV CLLLVTLALCCYQANAEFCPALVSELLDFFFI SEPLFKLSLAKFDAPPEAVAAKLGVKRCTDQ MSLQKRSLIAEVLVKILKKCSV
928	2278	A	8004	130	588	LAPLRCQPGTRTQPRSHPAANDPSAAMSAAG ARGLRATYHRLLDKVELMLPEKLRPLYNHPA GPRTVFFWAPIMKWGLVCAGLADMARPAEK LSTAQSAVLMATGFIWSRYSLVIIPKNWSLFA VNFFVGAAGASQLFRIWRYNQELKAKAHK
929	2279	A	8007	2	1016	EFARRRVFIAAREMSLLRSLRVFLVARTGSYP AGSLLRQSPQPRHTFYAGPRLSASASSKELLM KLRRKTGYSFVNCKKALETCGGDLKQAEIWL HKEAQKEGWSKAAKLQGRKTKEGLIGLLQE
030	2222					GNTTVLVEVNCETDFVSRNLKFQLLVQQVAL GTMMHCQTLKDQPSAYSKGFLNSSELSGLPA GPDREGSLKDQLALAIGKLGENMILKRAAWV KVPSGFYVGSYVHGAMQSPSLHKLVLGKYG ALVICETSEQKTNLEDVGRRLGQHVVGMAPL SVGSLDDEPGGEAETKMLSQPYLLDPSITLGQ YVQPQGVSVVDFVRFECGEGEEAAETE
930		A	8008	3	1679	NSRVWGPWTEPSAGSLRPMARKQNRNSKEL GLVPLTDDTSHAGPPGPGRALLECDHLRSGV PGGRRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYIKAFYNESWERRHGRPIDPDTL TLLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTTILTLQDHAPWVPYLSIVGILAILASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP
931	2281	A	8009	861	300	AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL AEMGLAVDPNRAVPLRKRKVKAMEVDIEER

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ĺ	USSN 09/496	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		914	correspondi	to last amino acid residue	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first amino acid	of peptide	Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
		ì		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	j	j]	peptide	sequence	/=possible nucleotide deletion, \-possible
	i		!	sequence		nucleotide insertion
		 	 	sequence		PKELVRKPYVLNDLEAEASLPEKKGNTLSRD
		1				LIDYVRYMVENHGEDYKAMARDEKNYYOD
ł		1	İ	ł	i	TPKQIRSKINVYKRFYPAEWQDFLDSLQKRK
						MEVE TRICTTAEWQDFEDSEQRAX
932	2282	Ā	8011	412	1	SNLCLGNSWRWRWAKSRHHCIPTVTLSKRSG
102	2202	1.	.0011] ''-	! *	DIRGSHFSSPQRQRSQRVPGKETARVLRAGK
		ĺ				QGRGQIPIPCPWPPPPPPPPPGSPGPGCRQFHQ
						SLEAKARHPASVREMRGKVKMRRALRRAPA
			-			STRASSROPNPK
933	2283	A	8012	147	1077	PPVPPASRSDMAQNLKDLAGRLPAGPRGMGT
1	2203		3312	*		ALKLLLGAGAVAYGVRESVFTVEGGHRAIFF
						NRIGGVQQDTILAEGLHFRIPWFQYPIIYDIRA
]			1		RPRKISSPTGSKDLQMVNISLRVLSRPNAQEL
						PSMYQRLGLDYEERVLPSIVNEVLKSVVAKF
						NASQLITQRAQVSLLIRRELTERAKDFSLILDD
						VAITELSFSREYTAAVEAKQVAQQEAQRAQF
						LVEKAKQEQRQKIVQAEGEAEAAKMLGEAL
]		1				SKNPGYIKLRKIRAAQNISKTIATSQNRIYLTA
					,	DNLVLNLQDESFTRGSDSLIKGKK
934 -	2284	A	8023	255	982	SQFSLSQVLVDSAEEGSLAAAAELAAQKREQ
73.		1	0020	233.	702	RLRKFRELHLMRNEARKLNHQEVVEEDKRL
l		Į				KLPANWEAKKARLEWELKEEEKKKECAARG
						EDYEKVKLLEISAEDAERWERKKKRKNPDLG
						FSDYAAAQLRQYHRLTKQIKPDMETYERLRE
			'			KHGEEFFPTSNSLLHGTHVPSTEEIDRMVIDLE
						KQIEKRDKYSRRRPYNDDADIDYINERNAKF
						NKKAERFYGKYTAEIKQNLERGTAV
935	2285	A	8027	59	310	LVSSTVNLLTEKAPWNSLAWTVTSYVFLKFL
		1				QGGGTGSTGMRDSALTLLGIGPSHRHSLSIRL
						SQHSSPAPMYSQTFHILVLG
936	2286	Α	8032	1	639	SGRECNMAKTYDYLFKLLLIGDSGVGKTCVL
						FRFSEDAFNSTFISTIGIDFKIRTIELDGKRIKLQ
		1				IWDTAGQERFRTITTAYYRGAMGIMLVYDIT
						NEKSFDNIRNWIRNIEEHASADVEKMILGNKC
			ļ			DVNDKRQVSKERGEKLALDYGIKFMETSAK
		ļ				ANINVENAFFTLARDIKAKMDKKLEGNSPQG
		}				SNQGVKITPDQQKRSSFFRCVLL
937	2287	A	8039	393	311	EETIHSENSYILEKYIPISANLTLTIA
938	2288	Α	8052	675	-1334	LHPAATSTAWLHVPPGLSMALSWVLTVLSLL
			ł			PLLEAQIPLCANLVPVPITNATLDRITGKWFYI
				l		ASAFRNEEYNKSVQEIQATFFYFTPNKTEDTIF
					ĺ	LREYQTRQDQCIYNTTYLNVQRENGTISRYV
						GGQEHFAHLLILRDTKTYMLAFDVNDEKNW
]		1	GLSVYADKPETTKEQLGEFYEALDCLRIPKSD
					1	VVYTDWKKDKCEPLEKQHEKERKQEEGES
939	2289	A	8055	12	1039	SSVAEFPERVQLSQPQNWNFSGAGGAWSLDF
					ļ	AEQLKWSAELARLGESIMDGKQGGMDGSKP
				ļ	ļ	AGPRDFPGIRLLSNPLMGDAVSDWSPMHEAA
						IHGHQLSLRNLISQGWAVNIITADHVSPLHEA
						CLGGHLSCVKILLKHGAQVNGVTADWHTPL
'				ļ		FNACVSGSWDCVNLLLOHGASVQPESDLASP
İ				1	ľ	IHEAARRGHVECVNSLIAYGGNIDHKISHLGT
						PLYLACENQQRACVKKLLESGADVNQGKGQ
						DSPLHAVARTASEELACLLMDFGADTQAKN
						AEGKRPVELVPPESPLAQLFLEREGPPSLMQL
						CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH
	ļ					L
940	2290	A	8058	2	1203	KVLSIREPAHSTARKASEPSQPSQPSQPGGHLI
						ARLRTMDLHLFDYSEPGNFSDISWPCNSSDCI

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion VVDTVMCPNMPNKSVLLYTLSFIYIFIFVIGMI ANSVVVWVNIQAKTTGYDTHCYILNLAIADL WVVLTIPVWVVSLVQHNQWPMGELTCKVTH LIFSINLFGSIFFLTCMSVDRYLSITYFTNTPSS
	-				·	RKKMVRRVVCILVWLLAFCVSLPDTYYLKT VTSASNNETYCRSFYPEHSIKEWLIGMELVSV VLGFAVPFSILAVFYFLLARAISASSDQEKHSS RKIIFSYVVVFLVCWLPYHVAVLLDIFSILHYI PFTCRLEHALFTALHVTQCLSLVHCCVNPVL YSFINRNYRYELMKAFIFKYSAKTGLTKLIDA SRVSETEYSALEQSTK
941	2291	A	8059	73	432	DMAGLMTIVTSLLFLGVCAHHIIPTGSVVLPS PCCMFFVSKRIPENRVVSYQLSSRSTCLKAGV IFTTKKGQQFCGDPKQEWVQRYMKNLDAKQ KKASPRARAVAVKGPVQRYPGNQTTC
942	2292	A	8067	278	1262	GGIGEIKQRPSCLGRCLDPSLSVLMNISLGLGS VFSAVISQKPSRDICQRGTSLTIQCQVDSQVT MMFWYRQQPGQSLTLIATANQGSEATYESGF VIDKFPISRPNLTFSTLTVSNMSPEDSSIYLCSA GRQGTYEQYFGPGTRLTVTEDLKNVFPPEVA VFEPSEAEISHTQKATLVCLATGFYPDHVELS WWVNGKEVHSGVSTDPQPLKEQPALNDSRY CLSSRLRVSATFWQNPRNHFRCQVQFYGLSE NDEWTQDRAKPVTQIVSAEAWGRADCGFTS ESYQQGVLSATILYEILLGKATLYAVLVSALV LMAMVKRKDSRG
943	2293	A	8070	1	879	MVKVVPATRGNLPRSQLTGTHQHCQPREPKI TASERLRRPRATARLRAHAAPPEPPLAVFAP PSDRKELLALPVACDPVIASVMSWVQAASLI QGPGDKGDVFDEEADESLLAQREWQSNMQR RVKEGYRDGIDAGKAVTLQQGFNQGYKKGA
				-		EVILNYGRLRGTLSALLSWCHLHNNNSTLINK INNLLDAVGQCEEYVLKHLKSITPPSHVVDLL DSIEDMDLCHVVPAEKKIDEAKDERLCENNA EFNKNCSKSHSGIDCSYVECCRTQEHAHSGK PKPHMDFGTDSQF
944	2294	А	8073	1	797 i	ESARWSRQLRRTLIRLSFPISCGRSHAFGGCK MAATSGTDEPVSGELVSVAHALSLPAESYGN DPDIEMAWAMRAMQHAEVYYKLISSVDPQF LKLTKVDDQIYSEFRKNFETLRIDVLDPEELK SESAKEK WRPFCLKFNGIVEDFNYGTLLRLD CSQGYTEENTIFAPRIQFFAIEIARNREGYNKA VYISVQDKEGEKGVNNGGEKRADSGEEENT KNGGEKGADSGEEKEEGINREDKTDKGGEK GKEADKEINKSGEKAM
945	2295	A .	8074	2	505	GAATLLRSASSAARKAAEAEQVWLHLHRYL SADRRVLGLREWGRPASERECSLCQRLKREL NMGDVEKGKKIFIMKCSQCHTVEKGGKHKT GPNLHGLFGRKTGQAPGYSYTAANKNKGIIW GEDTLMEYLENPKKYIPGTKMIFVGIKKKEER ADLIAYLKKATNE
946	2296	A	8081	42	590	EGRRGKFGGKLCNFLFYFHSNSAESRMDVLF VAIFAVPLILGQEYEDEERLGEDEYYQVVYY YTVTPSYDDFSADFTIDYSIFESEDRLNRLDK DITEAIETTISLETARADHPKPVTVKPVTTEPQ SPRSEAMPCPVLRSPIPLPPVRVPLFRWGCISC KKVGRRLLMTLWMGVWQEEIGR
947	2297	A	8084	322	549	GGGSSPRELAGAAGLTVISQAVAARRQQPSF SRARAPAHSLRAALSLASSARSWGAVSRDRG

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		i	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
· I		1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	ļ			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	Ì	i	İ	peptide	İ	/=possible nucleotide deletion, \=possible
			ļ	sequence		nucleotide insertion
948	2298	В	8093	2006	046	PCPPAIMYQSSNKC
940	2298	В	8093	3905	846	MEPGEVKDRILENISLSVKKLQSYFAACEDEI
				1		PAIRNHDKVLQRLCEHLDHALLYGLQDLSSG
)]		j]	YWVLVVHFTRREAIKQIEVLQHVATNLGRSR AWLYLALNENSLESYLRLFQENLGLLHKYYV
1]				KNALVCSHDHLTLFLTLVSGLEFIRFELDLDA
		1		i		PYLDLAPYMPDYYKPQYLLDFEDRLPSSVHG
1 1		l	}	i		SDSLSLNSFNSVTSTNLEWDDSAIAPSSEDYD
						FGDVFPAVPSVPSTDWEDGDLTDTVSGPRST
						ASDLTSSKASTRSPTQRQNPFNEEPAETVSSS
		ĺ	!			DTTPVHTTSQEKEEAQALDPPDACTELEVIRV
1			i			TKKKKIGKKKKSRSDEEASPLHPACSQKKCA
						KQGDGDSRNGSPSLGRDSPDTMLASPQEEGE
						GPSSTTESSERSEPGLLIPEMKDTSMERLGQPL
						SKVIDQLNGQLDPSTWCSRAEPPDQSFRTGSP
] }			j i			GDAPERPPLCDFSEGLSAPMDFYRFTVESPST
						VTSGGGHHDPAGLGQPLHVPSSPEAAGQEEE
						GGGGEGQTPRPLEDTTREAQELEAQLSLVRE
1 1			i			GPVSEPEPGTQEVLCQLKRDQPSPCLSSAEDS
						GVDEGQGSPSEMVHSSEFRVDNNHLLLLMIH VFRENEEQLFKMIRMSTGHMEGNLQLLYVLL
						TDCYVYLLRKGATEKPYLVEEAVSYNELDY
1 1	'		' '			VSVGLDQQTVKLVCTNRRKQFLLDTADVAL
						AEFFLASLKSAMIKGCREPPYPSILTDATMEK
1						LALAKFVAQESKCEASAVTVRFYGLVHWED
1 [PTDESLGPTPCHCSPPEGTITKEGMLHYKAGT
						SYLGKEHWKTCFVVLSNGILYQYPDRTDVIP
						LLSVNMGGEQCGGCRRANTTDRPHAFQVILS
1						DPPCLELSAESEAEMAEWMQHLCQAVSKGVI
1 1				+		PQGVAPSPCIPCCLVLTDDRLFTCHEDCQTSF
						FRSLGTAKLGDISAVSTEPGKEYCVLEFSQDS
1 1					İ	QQLLPPWVIYLSCTSELDRLLSALNSGWKTIY
! !						QVDLPHTAIQEASNKKKFEDALSLIHSAWQR
949	2299	A	8095	9	2374	SDSLCRGRASRDPWC*
1747	2277	^	0073	,	2374	ARRADTVLLESPSMLQGLLPVSLLLSVAVSAI KELPGVKKYEVVYPIRLHPLHKREAKEPEQQ
1	1	j		İ	j	EQFETELKYKMTINGKIAVLYLKKNKNLLAP
		[ĺ			GYTETYYNSTGKEITTSPQIMDDCYYQGHILN
		ļ				EKVSDASISTCRGLRGYFSQGDQRYFIEPLSPI
	ļ					HRDGQEHALFKYNPDEKNYDSTCGMDGVL
			ĺ			WAHDLQQNIALPATKLVKLKDRKVQEHEKY
	İ					IEYYLVLDNGEFKRYNENQDEIRKRVFEMAN
	ł	ŀ	Ì			YVNMLYKKLNTHVALVGMEIWTDKDKIKIT
		- 1				PNASFTLENFSKWRGSVLSRRKRHDIAQLITA
	İ	1				TELAGTTVGLAFMSTMCSPYSVGVVQDHSD
[ľ	ĺ	ľ			NLLRVAGTMAHEMGHNFGMFHDDYSCKCPS
	ļ				1	TICVMDKALSFYIPTDFSSCSRLSYDKFFEDKL
	j	J	J	J	ļ	SNCLFNAPLPTDIISTPICGNQLVEMGEDCDC GTSEECTNICCDAKTCKIKATFQCALGECCEK
	1			1	ŀ	CQFKKAGMVCRPAKDECDLPEMCNGKSGNC
	-	l				PDDRFQVNGFPCHHGKGHCLMGTCPTLOEQ
	- 1				ļ	CTELWGPGTEVADKSCYNRNEGGSKYGYCR
					l	RVDDTLIPCKANDTMCGKLFCQGGSDNLPW
		l	1		!	KGRIVTFLTCKTFDPEDTSQEIGMVANGTKCG
			ł	1		DNKVCINAECVDIEKAYKSTNCSSKCKGHAV
						CDHELQCQCEEGWIPPDCDDSSVVFHFSIVVG
]				İ		VLFPMAVIFVVVAMVIRHQSSREKQKKDQRP
	ļ	ĺ		1	ĺ	LSTTGTRPHKQKRKPQMVKAVQPQEMSQMK
			-			PHVYDLPVEGNEPPASFHKDTNALPPTVFKD
L					<u> </u>	NPMSTPKDSNPKA

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first	Predicted end nucleotide location corresponding to last amino acid residue	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
				amino acid residue of peptide sequence	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
950	2300	A	8100	1	1251	MGLLLMILASAVLGSFLTLLAQFFLLYRRQPE PPADEAARAGEGFRYIKPVPGLLLREYLYGG GRDEEPSGAAPEGGATPTAAPETPAPPTRETC YFLNATILFLFRELRDTALTRRWVTKKIKVEF EELLQTKTAGRLLEGLSLRDVFLGETVPFIKTI RLVRPVVPSATGEPDGPEGEALPAACPEELAF EAEVEYNGGFHLAIDVDLVFGKSAYLFVKLS RVVGRLRLVFTRVPFTHWFFSFVEDPLIDFEV RSQFEGRPMPQLTSIIVNQLKKIIKRKHTLPNY KIRFKPFFPYQTLQGFEEDEEHIHIQQWALTE GRLKVTLLECSRLLIFGSYDREANVHCTLELS SSVWEEKQRSSIKTGTISLTAVFMGWHRVSE AFPGLWYKLLVDLPFWGLEDGGPLLTVPLRQ CPG
951	2301	A	8108	1612	839	EVALFCFEMAAGMYLEHYLDSIENLPFELQR NFQLMRDLDQRTEDLKAEIDKLATEYMSSAR SLSSEEKLALLKQIQEAYGKCKEFGDDKVQL AMQTYEMVDKHIRRLDTDLARFEADLKEKQI ESSDYDSSSSKGKKGRTQKEKKAARARSKG KNSDEEAPKTAQKKLKLVRTSPEYGMPSVTF GSVHPSDVLDMPVDPNEPTYCLCHQVSYGE MIGCDNPDCSIEWFHFACVGLTTKPRGKWFC PRCSQERKKK
952	2302	A	8112	595	291	PSVASLARRFSGRALWPPSHSVPGNRALCPRL LHGTTLPGGNQRELARQKNMKKQSDSVKGK RRDDGLSAAARKQRDSTPRDSEIMQQKQKK ANEKKEEPK
953	2303	A	8118	1	669	VCAGIRDPCSTPLAKPAAGGAENLSFGKQPG LETNILKMTTPNKTPPGADPKQLERTGTVREI GSQAVWSLSSCKPGFGVDQLRDDNLETYWQ SDGSQPHLVNIQFRRKTTVKTLCIYADYKSDE
,						SYTPSKISVRVGNNFHNLQEIRQLELVEPSGW IHVPLTDNHKKPTRTFMIQIAVLANHQNGRD THMRQIKIYTPVEESSIGKFPRCTTIDFMMYRS IR
954	2304	A	8133	66	1015	PPLPPRSFPNLFSRPEPLPEPGRRGCNRSREPA ARAPSPPPPFEGAPGRAMVKVTFNSALAQKE AKKDEPKSGEEALIIPPDAVAVDCKDPDDVV PVGQRRAWCWCMCFGLAFMLAGVILGGAY LYKYFALQPDDVYPCGIKYIKDDVILNEPSAD APAALYQTIEENIKIFEEEVEFISVPVPEFADS DPANIVHDFNKKLTAYLDLNLDKCYVIPLNT SIVMPRNILLELLINIKAGTYLPQSYLIHEHMV ITDRIENIDHLGFFIYRLCHDKETYKLQRRETI KGIQKREASNCFAIRHFENKFAVETLICS
955	2305	A	8143	35	1171	VESRSAWHEGEDQIDRLDFIRNQMNLLTLDV KKKIKEVTEEVANKVSCAMTDEICRLSVLVD EFCSEFHPNPDVLKIYKSELNKHIEDGMGRNL ADRCTDEVNALVLQTQQEIIENLKPLLPAGIQ DKLHTLIPCKKFDLSYNLNYHKLCSDFQEDIV FRFSLGWSSLVHRFLGPRNAQRVLLGLSEPIF QLPRSLASTPTAPTTPATPDNASQEELMITLVT GLASVTSRTSMGIIIVGGVIWKTIGWKLLSVS. LTMYGALYLYERLSWTTHAKERAFKQQFVN YATEKLRMIVSSTSANCSHQVKQQIATTFARL CQQVDITQKQLEEIARLPKEIDQLEKIQNNS KLLRNKAVQLENELENFTKQFLPSSNEES
956	2306	A	8157	1854	798	ASGSPAPSSSSAMAAACGPGAAGYCLLLGLH LFLLTAGPALGWNDPDRMLLRDVKALTLHY

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						DRYTTSRRLDPIPQLKCVGGTAGCDSYTPKVI QCQNKGWDGYDVQWECKTDLDIAYKFGKT VVSCEGYESSEDQYVLRGSCGLEYNLDYTEL GLQKLKESGKQHGFASFSDYYYKWSSADSC NMSGLITIVVLLGIAFVVYKLFLSDGQYSPPP YSEYPPFSHRYQRFTNSAGPPPPGFKSEFTGPQ NTGHGATSGFGSAFTGQQGYENSGPGFWTGL GTGGILGYLFGSNRAATPFSDSWYYPSYPPSY PGTWNRAYSPLHGGSGSYSVCSNSDTKTRTA SGYGGTRRR
957	2307	A	8159	1492	528	THVVMTGMCYAPHQVLSYINGVTTSKPGVSL VYSMPSRNLSLRLEGLQEKDSGPYSCSVNVQ DKQGKSRGHSIKTLELNVLVPPAPPSCRLQGV PHVGANVTLSCQSPRSKPAVQYQWDRQLPSF QTFFAPALDVIRGSLSLTNLSSSMAGVYVCKA HNEVGTAQCNVTLEVSTGPGAAVVAGAVVG TLVGLGLLAGLVLLYHRRGKALEEPANDIKE DAIAPRTLPWPKSSDTISKNGTLSSVTSARAL RPPHGPPRPGALTPTPSLSSQALPSPRLPTTDG AHPQPISPIPGGVSSSGLSRMGAVPVMVPAQS QAGSLV
958	2308	A	8161	2340	1192	ELARRPKQQSSEKSRNMIRNWLTIFILFPLKLV EKCESSVSLTVPPVVKLENGSSTNVSLTLRPP LNATLVITFEITFRSKNITILELPDEVVVPPGVT NSSFQVTSQNVGQLTVYLHGNHSNQTGPRIR FLVIRSSAISINQVIGWIYFVAWSISFYPQVIM NWRRKSVIGLSFDFVALNLTGFVAYSVFNIGL LWVPYIKEQFLLKYPNGVNPVNSNDVFFSLH AVVLTLIIIVQCCLYERGGQRVSWPAIGFLVL AWLFAFVTMIVAAVGVITWLQFLFCFSYIKL AVTLVKYFPQAYMNFYYKSTEGWSIGNVLL DFTGGSFSLLQMFLQSYNNDQWTLIFGDPTK FGLGVFSIVFDVVFFIQHFCLYRKRPGYDQLN
959	2309	Α	8163	521	1345	GERAGRRGRLGVWAQPQPLLPRPVGSRRE MQPPGPPPAYAPTNGDFTFVSSADAEDLSGSI ASPDVKLNLGGDFIKESTATTFLRQRGYGWL LEVEDDDPEDNKPLLEELDIDLKDIYYKIRCV LMPMPSLGFNRQVVRDNPDF WGPLAVVLFFS MISLYGQFRVVSWIITIWIFGSLTIFLLARVLG GEVAYGQVLGVIGYSLLPLIVIAPVLLVVGSF EVVSTLIKLFGVFWAAYSAASLLVGEEFKTK KPLLIYPIFLLYIYFLSLYTGV
960	2310	A	8167		2921	MTCFKGQKGEQRSHAFEANKDHKAKVPSPN LYSQLNALQFTVDERSILWLNQFLLDLKQSL NQFMAVYKLNDNSKSDEHVDVRVDGLMLK FVIPSEVKSECHQDQPRAISIQSSEMIATNTRH CPNCRHSDLEALFQDFKDCDFFSKTYTSFPKS CDNFNLLHPIFQRHAHEQDTKMHEIYKGNITP QLNKNTLKTSAATDVWAVYFSQFWIDYEGM KSGKGRPISFVDSFPLSIWICQPTRYAESQKEP QTCNQVSLNTSQSESSDLAGRLKRKKLLKEY YSTESEPLTNGGQKPSSSDTFFRFSPSSSEADI HLLVHVHKHVSMQINHYQYLLLLFLHESLILL SENLRKDVEAVTGSPASQTSICIGILLRSAELA LLLHPVDQANTLKSPVSESVSPVVPDYLPTEN GDFLSSKRKQISRDINRIRSVTVNHMSDNRSM SVDLSHIPLKDPLLFKSASDTNLQKGISFMDY LSDKHLGKISEDESSGLVYKSGSGEIGSETSD KKDSFYTDSSSVLNYREDSNILSFDSDGNQNI LSSTLTSKGNETIESIFKAEDLLPEAASLSENL

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valline, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion DISKEETPPVRTLKSQSSLSGKPKERCPPNLAP LCVSYKNMKRSSSQMSLDTISLDSMILEEQLL ESDGSDSHMFLEKGNKKNSTTNYRGTAESVN AGANLQNYGETSPDAISTNSEGAQENHDDLM SVVVFKITGVNGEIDIRGEDTEICLQVNQVTP DQLGNISLRHYLCNRPVGSDQKAVIHSKSSPE ISLRFESGPGAVIHSLLAEKNGFLQCHIENFST EFLTSSLMNIQHFLEDETVATVMPMKIQVSNTKINLKDDSPRSSTVSLEPAPVTVHIDHLVVER SDDGSFHIRDSHMLNTGNDLKENVKSDSVLL TSGKYDLKKQRSVTQATQTSPGVPWPSQSAN FPEFSFDFTREQLMEENESLKQELAKAKMAL
961	2311	A	8172	1442	682	AEAHLEKDALLHHIKKMTVE TAAMSIFTPTNQIRLTNVAVVRMKRAGKRFEI ACYKNKVVGWRSGVEKDLDEVLQTHSVFVN VSKGQVAKKEDLISAFGTDDQTEICKQILTKG EVQVSDKERHTQLEQMFRDIATIVADKCVNP ETKRPYTVILIERAMKDIHYSVKTNKSTKQQA LEVIKQLKEKMKIERAHMRLRFILPVNEGKKL KEKLKPLIKVIESEDYGQQLEIVCLIDPGCFREI DELIKKETKGKGSLEVLNLKDVEEGDEKFE
962	2312	A	8175	286	587	NISNKAEVSSHPSVISHSMDSFGQPRPEDNQS VLRRMQKKYWKTKQVFIKATGKKEDEHLVA SDAELDAKLEVFHSVQETCTELLKIIEKYQLR
963	2313	Ą	8181	13	2215	LNGMKS AEGCAERRGTEPVVELSMSWESGAGPGLGSQ GMDLVWSAWYGKCVKGKGSLPLSAHGIVV AWLSRAEWDQVTVYLFCDDHKLQRYALNRI TVWRSRSGNELPLAVASTADLIRCKLLDVTG GLGTDELRLLYGMALVRFVNLISERKTKFAK VPLKCLAQEVNIPDWIVDLRHELTHKKMPHI
						NDCRRGCYFVLDWLQKTYWCRQLENSLRET WELEEFREGIEEEDQEEDKNIVVDDITEQKPE PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT GRNARRFSAGQWEARRGWRLFNCSASLDWP RMVESCLGSPCWASPQLLRIIFKAMGQGLPD EEQEKLLRICSIYTQSGENSLVQEGSEASPIGK SPYTLDSLYWSVKPASSSFGSEAKAQQQEEQ GSVNDVKEEEKEEKEVLPDQVEEEEENDDQE EEEEDEDDEDDEEDRMEVGPFSTGQESPTA ENARLLAQKRGALQGSAWQVSSEDVRWDTF PLGRMPGQTEDPAELMLENYDTMYLLDQPV LEQRLEPSTCKTDTLGLSCGVGSGNCSNSSSS NFEGLLWSQGQLHGLKTGLQLF
964	2314	A	8184	6	1393	EPRNFRDDSTRPRTRGRTRGRRRRACRSAE GTGLRSLLLPPRLQLPAGPFSRCRWDPVSSRR PSTMPPKKGGDGIKPPPIIGRFGTSLKIGIVGLP NVGKSTFFNVLTNSQASAENFPFCTIDPNESR VPVPDERFDFLCQYHKPASKIPAFLNVVDIAG LVKGAHNGQGLGNAFLSHISACDGIFHLTRA FEDDDITHVEGSVDPIRDIEIIHEELQLKDEEMI GPIIDKLEKVAVRGGDKKLKPEYDIMCKVKS WVIDQKKPVRFYHDWNDKEIEVLNKHLFLTS KPMVYLVNLSEKDYIRKKNKWLIKIKEWVD KYDPGALVIPFSGALELKLQELSAEERQKYLE

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
1	peptide		in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
	seq-	İ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
1	uence	[09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
			-	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			ļ	peptide		/-possible nucleotide deletion, \-possible
				sequence	ļ	nucleotide insertion
						ANMTQSALPKIIKAGFAALQLEYFFTAGPDEV
						RAWTIRKGTKAPQAAGKIHTDFEKGFIMAEV
1						MKYEDFKEEGSENAVKAAGKYRQQGRNYIV
L						EDGDIIFFKFNTPQQPKKK
965	2315	A	8195	1437	594	RSFSLSFSLLSPSEMMALGAAGATRVFVAMV
1 1					}	AAALGGHPLLGVSATLNSVLNSNAIKNLPPPL
						GGAAGHPGSAVSAAPGILYPGGNKYQTIDNY
						QPYPCAEDEECGTDEYCASPTRGGDAGVQIC
						LACRKRRKRCMRHAMCCPGNYCKNGICVSS
					!	DQNHFRGEIEETITESFGNDHSTLDGYSRRTT
						LSSKMYHTKGQEGSVCLRSSDCASGLCCARH
1						FWSKICKPVLKEGQVCTKHRRKGSHGLEIFQ
066	2216		0207	416	4082	RCYCGEGLSCRIQKDHHQASNSSRLHTCQRH
966	2316	A	8207	410	4002	KFKLIKIMLLTLIILLPVVSKFSFVSLSAPQHW SCPEGTLAGNGNSTCVGPAPFLIFSHGNSIFRI
						DTEGTNYEOLVVDAGVSVIMDFHYNEKRIY
1]						WVDLEROLLORVFLNGSROERVCNIEKNVSG
1 1						MAINWINEEVIWSNQQEGIITVTDMKGNNSHI
1					-	LLSALKYPANVAVDPVERFIFWSSEVAGSLY
1						RADLDGVGVKALLETSEKITAVSLDVLDKRL
1						FWIQYNREGSNSLICSCDYDGGSVHISKHPTQ
						HNLFAMSLFGDRIFYSTWKMKTIWIANKHTG
						KDMVRINLHSSFVPLGELKVVHPLAQPKAED
						DTWEPEQKLCKLRKGNCSSTVCGQDLQSHLC
			1			MCAEGYALSRDRKYCEGNDWKYCEDVNEC
1 1					.	AFWNHGCTLGCKNTPGSYYCTCPVGFVLLPD
[GKRCHQLVSCPRNVSECSHDCVLTSEGPLCF
						CPEGSVLERDGKTCSGCSSPDNGGCSQLCVPL
1 1						SPVSWECDCFPGYDLQLDEKSCAASGPQPFL
1						LFANSQDIRHMHFDGTDYGTLLSQQMGMVY
						ALDHDPVENKIYFAHTALKWIERANMDGSQ
						RERLIEEGVDVPEGLAVDWIGRRFYWTDRGK
						SLIGRSDLNGKRSKIITIENISQPRGIAVHPMAK
1 1						RLFWTDTGINPRIESSSLQGLGRLVIASSDLIW
						PSGITIDFLTDKLYWCDAKQSVIEMANLDGSK
						RRRLTQNDVGHPFAVAVFEDYVWFSDWAMP SVIRVNKRTGKDRVRLQGSMLKPSSLVVVHP
						LAKPGADPCLYONGGCEHICKKRLGTAWCS
						CREGFMKASDGKTCLALDGHQLLAGGEVDL
						KNOVTPLDILSKTRVSEDNITESQHMLVAEIM
1 .						VSDQDDCAPVGCSMYARCISEGEDATCQCLK
					1	GFAGDGKLCSDIDECEMGVPVCPPASSKCINT
	ļ					EGGYVCRCSEGYQGDGIHCLDIDECQLGVHS
]		CGENASCTNTEGGYTCMCAGRLSEPGLICPD
				j		STPPPHLREDDHHYSVRNSDSECPLSHDGYCL
			l		!	HDGVCMYIEALDKYACNCVVGYIGERCQYR
				1		DLKWWELRHAGHGQQQKVIVVAVCVVVLV
	,					MLLLLSLWGAHYYRTQKLLSKNPKNPYEESS
1		[ĺ		1	RDVRSRRPADTEDGMSSCPQPWFVVIKEHQD
						LKNGGQPVAGEDGQAADGSMQPTSWRQEPQ
						LCGMGTEQGCWIPVSSDKGSCPQVMERSFH
1	1			İ		MPSYGTQTLEGGVEKPHSLLSANPLWQQRAL
						DPPHQMELTQ
967	2317	Α	8210	3	601	SSAMGSRSSHAAVIPDGDSIRRETGFSQASLL
]	j	l J		ļ		RLHHRFRALDRNKKGYLSRMDLQQIGALAV
						NPLGDRIIESFFPDGSQRVDFPGFVRVLAHFRP
	.					VEDEDTETQDPKKPEPLNSRRNKLHYAFQLY
	·					VEDEDTETQDPKKPEPLNSRRNKLHYAFQLY DLDRDGKISRHEMLQVLRLMVGVQVTEEQL
						VEDEDTETQDPKKPEPLNSRRNKLHYAFQLY

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
968	2318	A	8211	2	409	ISSCPHTAYEGSMSTLSNFTQTLEDVFRRIFIT YMDNWRQNTTAEQEALQAKVDAENFYYVIL YLMVMIGMFSFIIVAILVSTVKSKRREHSNDP YHQYIVEDWQEKYKSQILNLEESKATIHENIG AAGFKMSP
969	2319	A	8215			GMPRSRGGRAAPGPPPPPPPPGQAPRWSRWR VPGRLLLLLPALCCLPGAARAAAAAAGGN RAAVAVAVARADEAEAPFAGQNWLKSYGY LLPYDSRASALHSAKALQSAVSTMQQFYGIP VTGVLDQTTIEWMKKPRCGVPDHPHLSRRRR NKRYALTGQKWRQKHITYSIHNYTPKVGELD TRKAIRQAFDVWQKVTPLTFEEVPYHEIKSDR KEADIMIFFASGFHGDSSPFDGEGGFLAHAYF PGPGIGGDTHFDSDEPWTLGNANHDGNDLFL VAVHELGHALGLEHSSDPSAIMAPFYQYMET HNFKLPQDDLQGIQKIYGPPAEPLEPTRPLPTL PVRRIHSPSERKHERQPRPPRPPLGDRPSTPGT KPNICDGNFNTVALFRGEMFVFKDRWFWRL RNNRVQEGYPMQIEQFWKGLPARIDAAYER ADGRFVFFKGDKYWVFKEVTVEPGYPHSLG ELGSCLPREGIDTALRWEPVGKTYFFKGERY WRYSEERRATDPGYPKPITVWKGIPQAPQGA FISKEGYYTYFYKGRDYWKFDNQKLSVEPGY PRNILRDWMGCNQKEVERRKERRLPQDDVDI MVTINDVPGSVNAVAVVIPCILSLCILVLVYTI FQFKNKTGPQPVTYYKRPVQEWV
970	2320	A	8216	1235	2223	SRLSLQFYVSFRRTGLFTCKLIVEIFFRNYMN DSLRTNVFVRFQPETIACACIYLAARALQIPLP TRPHWFLLFGTTEEEIQEICIETLRLYTRKKPN YELLEKEVEKRKVALQEAKLKAKGLNPDGTP ALSTLGGFSPASKPSSPREVKAEEKSPISINVK
						TVKKEPEDRQQASKSPYNGVRKDSKRSRNSR SASRSRSRTRSRSRSHTPRRHYNNRRSRSGTY SSRSRSRSSHSESPRRHHNHGSPHLKAKHTR DDLKSSNRHGHKRKKSRSRSQSKSRDHSDAA KKHRHERGHHRDRRERSRSFERSHKSKHHGG SRSGHGRHRR
971	2321	A	8217	3	3274	DCRLQAAMPTNFTVVPVEAHADGGGDETAE RTEAPGTPEGPEPERPSPGDGNPRENSPFLNN VEVEQESFFEGKNMALFEEEMDSNPMVSSLL NKLANYTNLSQGVVEHEEDEESRREAKAPR MGTFIGVYLPCLQNILGVILFLRLTWIVGVAG VLESFLIVAMCCTCTMLTAISMSAIATNGVVP AGGSYYMISRSLGPEFGGAVGLCFYLGTTFA GAMYILGTIEIFLTYISPGAAIFQAEAAGGEAA AMLHNMRVYGTCTLVLMALVVFVGVKYVN KLALVFLACVVLSILAIYAGVIKSAFDPPDIPV CLLGNRTLSRRSFDACVKAYGIHNNSATSAL WGLFCNGSQPSAACDEYFIQNNVTEIQGIPGA ASGVFLENLWSTYAHAGAFVEKKGVPSVPV AEESRASTLPYVLTDIAASFTLLVGIYFPSVTG IMAGSNRSGDLKDAQKSIPTGTILAIVTTSFIY LSCIVLFGACIEGVVLRDKFGEALQGNLVIGM LAWPSPWVIVIGSFSTCGAGLQTLTGAPRLL QAIARDGIVPFLQVFGHGKANGEPTWALLLT VLICETGILIASLDSVAPILSMFFLMCYLFVNL ACAVQTLLRTPNWRPRFKFYHWTLSFLGMSL CLALMFICSWYYALSAMLIAGCIYKYIEYRG AEKEWGDGIRGLSLNAARYALLRVEHGPPHT KNWRPQVLVMLNLDAEQAMKHPRLLSFTSQ

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LKAGKGLTIVGSVLEGTYLDKHMEAQRAEE NIRSLMSTEKTKGFCQLVVSSSLRDGMSHLIQ SAGLGGLKHNTVLMA WPASWKQEDNPFSW KNFVDTVRDTTAAHQALLVAKNVDSFPQNQ ERFGGGHIDVWWIVHDGGMLMLLPFLLRQH KVWRKCRMRIFTVAQVDDNSIQMKKDLQMF LYHLRISAEVEVVEMVENDISAFTYERTLMM EQRSQMLKQMQLSKNEQEREAQLIHDRNTAS HTAAAARTQAPPTPDKVQMTWTREKLIAEK YRSRDTSLSGFKDLFSMKPDQSNVRRMHTAV KLNGVVLNKSQDAQLVLLNMPGPPKNRQGD ENYMEFLEVLTEGLNRVLLVRGGGREVTTIYS
972	2322	A	8224	701	246	TSRRVTMKFNPFVTSDRSKNRKRHFNAPSHV RRKIMSSPLSKELRQKYNVRSMPIRKDDEVQ VVRGHYKGQQIGKVVQVYRKKYVIYIERVQ REKANGTTVHVGIHPSKVVITRLKLDKDRKKI
973	2323	A	8237	279	4610	LERKAKSRQVGKEKGKYKEELIEKMQE GCPHAGGKGRVPTGGLTGGRTWSPSAAPRSC PRPGPTPAPGAMDKLPPSMRKRLYSLPQQVG AKAWIMDEEDAEEEGAGGRQDPSRRSIRLR PLPSPSPSAAAGGTESRSSALGAADSEGPARG AGKSSTNGDCRFRGSLASLGSRGGGSGGTG SGSSHGHLHDSAEERRLIAEGDASPGEDRTPP GLAAEPERPGASAQPAASPPPPQQPPQPASAS CEQPSVDTAIKVEGGAAAGDQILPEAEVRLG QAGFMQRQFGAMLQPGVNKFSLRMFGSQKA VEREQERVKSAGFWIIHPYSDFRFYWDLTML LLMVGNLIIIPVGITFFKDENTTPWIVFNVVSD TFFLIDLVLNFRTGIVVEDNTEIILDPQRIKMK YLKSWFMVDFISSIPVDYIFLIVETRIDSEVYK TARALRIVRFTKILSLLRLLRLSRLIRYIHQWE EIFHMTYDLASAVVRIVNLIGMMLLCHWDG CLQFLVPMLQDFPDDCWVSINNMVNNSWGK QYSYALFKAMSHMLCIGYGRQAPVGMSDV WLTMLSMIVGATCYAMFIGHATALIQSLDSS RRQYQEKYKQVEQYMSFHKLPPDTRQRIHD YYEHRYQGKMFDEESILGELSEPLREEIINFNC RKLVASMPLFANADPNFVTSMLTKLRFEVFQ PGDYIIREGTIGKKMYFIQHGVVSVLTKGNKE TKLADGSYFGEICLLTRGRRTASVRADTYCR LYSLSVDNFNEVLEEYPMMRAFETVALDRL DRIGKKNSILLHKVQHDLNSGVFNYQENEIIQ QIVQHDREMAHCAHRVQAAASATPTPTPVIW TPLIQAPLQAAAATTSVAIALTHHPRLPAAIFR PPPGSGLGNLGAGQTPRHLKRLQSLIPSALGS ASPASSPSQVDTPSSSSFHIQQLAGFSAPAGLS PLLPSSSSSPPPGACGSPSAPTPSAGVAATTIA GFGHFHKALGGSLSSSDSPLLTPLQPGARSPQ AAQPSPAPPGARGGLGPEHFLPPPPSSRSPSS SPGQLGQPPGELSLGLATGPLSTPETPPRQPEP PSLVAGASGGASPVGFTPRGGLSPPGHSSPGPP RTFPSAPPRASGSHGSLLLPPASSPPPPQVPQR RGTPPLTPGRLTQDLKLISASQPALPQDGAQT LRRASPHSSGESMAAFPLFPRAGGGSGGSGSS GGLGPPGRPYGAIPGQHVTLPRKTSSGSLPPP LSLFGARATSSGGPPLTAGPQREPGARPEPVR SKLPSNL EYKOWERFELSCONRNDIGYGKPRKGGGLI
974	2324	A	8247	279	468	EYKQWERRFLSCQNRNDLGYGKPRKGGGLL LVPVKDASRICSLTYLLGSHWNNLVVRSPVL

NO: of nucl- ouence uence USN seq- uence uence USN seq- uence uence USN seq- uence uence uence usy to the sequence use to the sequence uence uence uence uence usy to the sequence uence usy to the sequence peptide amino acid residue of peptide sequence peptide sequence sequence sequence sequence sequence use to the sequence per to the sequence per to the sequence per to the sequence sequence sequence sequence sequence sequence per to the sequence per to the sequence s	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucle coilde coilde coilde sequence			1				D=Aspartic Acid. E=Glutamic Acid
eotide sequence ue						1	F=Phenylalanine, G=Glycine H=Histidine
seq. uence 914 914 orrespondi plat in sat amino acid recidue of peptide sequence of peptide sequence of peptide sequence of peptide sequence of peptide sequence of peptide sequence of peptide sequence of peptide sequence of peptide sequence of peptide sequence of peptide sequence of peptide sequence of peptide sequence of peptide sequence of peptide sequence of peptide of	eotide		1	USSN			I=Isoleucine, K=Lvsine, L=Leucine
uence 914 and to first a min asaid residue of peptide sequence peptide sequence 1571 177	seq-	uence	ĺ	09/496	correspondi		
amino acid recidue of peptide requence very prossin, x 1—Natione, x—Sugine, very peodon, y—possible nucleotide deletion, y—possible nucleotide deletion, y—possible nucleotide deletion, y—possible nucleotide deletion, y—possible nucleotide insertion with the property of	uence		!	914		1	
975 2325 A 8249 62 1571 LVALKNWKPKGTNIPAPQSPVFGEAVSGV MICHORITORIO POSSIBLE MICHORIO del eletion, \timespace possible muclostide insertion 1571 LVALKNWKPKGTNIPAPQSPVFGEAVSGV MTKVLGMAPVLGPRPPQEQVGFLMVKVE EEKGKYLPSLEMFRQERRGFGYHDTPGFR LSQLKVLCCEWLRPEHTKEQILELU-LEQ LSQLKVLCCEWLRPEHTKEQILELU-LEQ LSQLKVLCCEWLRPEHTKEQILELU-LEQ LSQLKVLCCEWLRPEHTKEQILELU-LEQ LSQLKVLCCEWLRPEHTKEQILELU-LEQ LSQLKVLCCEWLRPEHTKEQILELU-LEQ LSQLKVLCCEWLRPEHTKEQILELU-LEQ LSQLKVLCCEWLRPEHTKEQILELU-LEQ LSQLKVLCCEWLRPEHTKEQILELU-LEQ LSQLKVLCCEWLRPEHTKEQILELU-LEQ LSQLKVLCCEWLRPEHTKEQILEU-LYEQ LSQLKVLCCEWLRPEHTKEQILEU-LYEQ LSQLKVLCCEWLRPEHTKEQILEU-LYEQ LSQLKVLCEWLRPHKYESWGPLYIOESGES-GER DPRKYKDCRLSTOHESADERGKGAS-GAG- MSSNITKHRRTHTGEKPYVCTK-GACKA-FS- NSSNITKHRRTHTGEKPYVCTK-GACKA-FS- NSSNITKHRRTHTGEKPYVCTK-GACKA-FS- NSSNITKHRRTHTGEKPYVCTK-GACKA-FS- NSSNITKHRRTHTGEKPYVCTK-GACKA-FS- NSSNITKHRRTHTGEKPYVCTK-GACKA-FS- NSSNITKHRRTHTGEKPYVCTK-GACKA-FS- NSSNITKHRRTHTGEKPYVCTK-GACKA-FS- NSSNITKHRRTHTGEK-PYVCTK-GACKA-FS- NSSN							
Peptide sequence uncloside deletion, \(\to \) possible nucloside insertion 975 2325 A 8249 62 1571	1		1				Y=Tyrosine, X=Unknown *=Stop codon
975 2325 A 8249 62 1571 LVALKNWKPKGTNIPAPQSPVFGEAVSGV MTKVLGMAPVLGPRPPQEQVGPLMVKP EEKGKYLPSLEAMFRQRRPGGYBOTFGER LSQLRVLCCEWLRPEIHTKEQILELLVLEQ ILIPQELQAWVQBHCPESAEEAVTLLEDI DEPGHQVSTPPNEQKPVWERISSSGTAKE SMQRQPLETSHKYESWGPLYIQESGEEGE DPRKVRDCRLSTQHEESABEQKGSEAEGL DIISVILANKPEASLERQCVNLENEKGTKPP EAGSKKGRESVPTKFTPGERYTCAECGK NSSNLTKHRRTHTGERPYVCTKCGKAFGSSD HQRMHTEEAPYQCKDGCKAFSGKGSLIR HTGERPYQCKDGCKAFSGKGSLIR HTGERPYQCKDGCKAFSGKGSLIR HTGERPYQCKDGCKAFSGKGSLIR HTGERPYQCKDGCKAFSGKGSLIR HTGERPYQCKDGCKAFSGKGSLIR HTGERPYQCHDCKCKSTSGNAKE GERPYKCKEGGKAFNISSNTNKHHRIHTG PVWCHHCKGKTFCSKSNLSKHQRVHTGGE GFPYKCKEGGKAFNISSNTNKHHRIHTG PVMCHHCKGKTFCSKSNLSKHQRVHTGGF GFYYNLSLPKSTYDKMKARAV GERPYKCKEGGKAFNISSNTNKHHRIHTG PVMCHHCKGKTFCSKSNLSKHQRVHTGGF GFYYNLSLPKSTYDKMKARAV GERPYCKCKEGKAFNISSNTNKHRIHTISTSP KELAEATKTLLHSLGTLAGELFSMESSND QEVMFLTNVNSSSSTOTQAVSRIVCGGP GGLKIKSLNWYEDNNYKALFGGNGTEG GGLKIKSLNWYEDNNYKALFGGNGTEG GGLKIKSLNWYEDNNYKALFGGNGTEG TFYDNSTTPYCNDLMKNLESSLSRIIWKA PLLVGKLLYTPDTPATRQWAAEVNKTFOEI VFHDLEGMWELSPKIWTFMENSGELDI KFWAGIVFTGTTGSSELPHHVKYKRMCLDI VERTNKINGDYDPOPRADPFEDMRYVW FAYLQDVEQAIRVLTGTEKKTGVYMQ VFCYDDIFIKLEPTATEVUKINSMELLDI VERTNKINGDYDPOPRADPFEDMRYVW FAYLQDVVEQAIRVLTGTEKKTGVYMQ VFCYDDIFIKLEPTATEVUKINSMELLDI VERTNKINGDYDPOPRADPFEDMRYVW FAYLQDVVEQAIRVLTGTEKKTGVYMQ VFCYDDIFIKLEPTATEVUKINSMELLDI VERTNKINGDYDPOPRADPFEDMRYVW FAYLQDVVEQAIRVLTGTEKKTGVYMQ VFCYDDIFIKLEPTATEVUKINSMELDI VERTNKINGDYDPOPRADPFEDMRYVW FAYLQDVVEQAIRVLTGTEKKTGVYMQ VFCYDDIFIKLEPTATEVUKINSMELDI VERTNKINGDYDPOPRADPFEDMRYVW FAYLQCVENNUKLEPTATEVUKINSMELDI VERTNKINGDYDPOPRADPFEDMRYVW FAYLQCVENNUKLEPTATEVUKINSMELDI VERTNKINGDYDPOPRADPFEDMRYVW FAYLQCVENNUKLEPTATEVUKINSMELDI VERTNKINGDYDPOPRADPFEDMRYVW FAYLQCVENNUKLEPTATEVUKINSMELDI VERTNKINGDYDPOPRADPFEDMRYVW FAYLQCVENNUKLEPTATEVUKINSMELDI VERTNKINGDYDPOPRADPFEDMRYVW FAYLQCVENNUKLEPTATEVUKINGSMELDI VERTNKINGDYDPOPRADFEDMRYVW FAYLQCVENNUKLEPTATEVUKINGSMELDI VERTNKINGDYDPOPRADFICHTANOV VSRGGWELLLKYQGGTILLSTHEINDEAD GBRIAISHGKLCCVGSSLFIKHGIGGTGYYI GMCXLSVALAPVGGSKVULDEPTAGV VSRGGWELLLKYAQGTIILSTHEINDEAD GGRAL			Ī		peptide		/=possible nucleotide deletion \=possible
975 2325 A 8249 62 1571 IVALKNWKPKGTNIPAPQSPVPGEAVSGV EEKGKYLPSLEMFRQRRQFGYHOTFGRR LSQLRVLCCEWLRPEHTKEQUELULUE URPELQAWVQEHCPESABEAVTLLEDLEI DEFGHQVSTPPNEQKPYWEKISSSGTAKE SMQPQHETSHKYESWGHLVIGESGEEQEI DPRKVRDCRLSTQHEESADEQKGSEAGEI DISVIIANKPASLERQCVNLENEKGTKP EAGSKKGRESVPTKPTGERRYICAEGGX. NSSNLTKHRRHTTGEKPYVCKCGKAFSGKGSLIR- HITGEKPYQCNEGGKSFSQHAGLSSHQRLI- GEKPYKCKEGGKAFHSSNFNKHHRHTG PYWCHHGGKTTCSKSNLSKLIQPHUDNET GFKYKCKEGGKAFHSSNFNKHHRHTG PYWCHHGKTTCSKSNLSKLIQPHUDNET GFKYKCKEGGKAFHSSNFNKHHRHTG PYWCHHGKTTCSKSNLSKLIQPHUDNET GFKYKNSIVARLFSDARRLLLYSQKDTSM MRKVLRTLQGICKSSSNLKLQDFLUDNET GFLYHNLSLPKSTYDKMLRADVILHKYPL YQLHLTSLCNGKSSSENIQLGDGVSSLC REKLAAAERVLRSNMDILKPILRTLNSTSP KELAAATKTLLHSLGTLAGELFSMRSWC GGKKKSLNWYEDNNYKALFGGNGTEG GGKKIKSLNWYEDNNYKALFGGNGTEG TPYDNSTTPYCNDLMKNLESPLARTENDOAR KHEPDVQSNGSYTWREAFNETNOAR KHEPDVQSNGSYTWREAFNETNOAR KHEPDVQSNGSYTWREAFNETNOAR KHEPDVGSNGSTROAR KHEPDVGSNGSTR			}		• •		nucleotide insertion
MTKVLGMAPVLGPRPPQEOVGPLMYKVE EEKGKYLPSLEMFRQRRPG/DIDTGPR LSQLRVLCCEWLRPEHITKEQILELLULED ILPQELQAWVQEHCPESAEEAVTLLEDLEI ILPQELQAWVQEHCPESAEEAVTLLEDLEI ILPQELQAWVQEHCPESAEEAVTLLEDLEI DEPGHQVSTPPNEQKPWERISSGTAKE SMQOPI-ETSHKYESWGPLYIQESGEEGEI DPRKVRDCRLSTQHEESABEAGKSEAEGL DIISVILANKPEASLERQCVNLENEKGTKEPP EAGSKKGRESVPTKPTGERGIK/GAECGK. NSSNLTKHRRTHTGEKPYVCTKCGKAFGQSSDI HORMHTEEAPYQCNEGKAFSQHAGLSSHQRI GEKPYKCKEGGKAFNGSSNINSKHERHTG PYWCHHCGKTYCKGCKAFGQSSDI HORMHTEEAPYQCNEGKAFSQHAGLSSHQRI GEKPYKCKEGGKAFNHSSNINSKHERHTG PYWCHHCGKTYCSKSNLSKHQRYHTGEGI PWQCHILLWKNLTFRRRGTCOTA ONTHASIVARI PSDARRLLLYSQKDTSMI MRKVLRTLQQIKKSSSNLKLQDFLVDNETI GFLYHNLSLPKSTYDKMLRADVILHKYUL VAWPLFFILLISVRLSYPPYCHECHFPMK MRKVLRTLQQIKKSSSNLKLQDFLVDNETI GFLYHNLSLPKSTYDKMLRADVILHKYUL VQLHLTSLCNGSKSEEMIQLGDQEVSELCC REKLAAAERVLRSNMDILKPILRTILNSTSP KELAEATKTLHSLGTLAGELFSMRSWSD QEVMFLTNVNSSSSSTQIYQAVSRIVCGI GGLKIKSLNWYEDNNYKALFGGNGTEED TFYDNSTTPYCNDLMKNLESSELSRIIWKA PLLVGKILTYPDTPATRQWARNYTFOEI VFHDLEGMWEELSPKIWTFMENSQEMDLA MILDSRNNDHFWEQQLDGLDWTGDIDT VFHDLEGMWEELSPKIWTFMENSQEMDLA MILDSRNNDHFWEQQLDGLDWTGDIDT VFHDLEGMWEELSPKIWTFMENSQEMDLA KHEEDVQSNOSYTYWEAFANETNOAIR KPMECVNLNKLEPIATEVWLNKSMELL VSAGGLFOVALIRVLTGFEKKTGVYMQQ YPCYVDDIFLRVMSSNSSYTWEAFANETNOAIR KFMECVNLNKLEPIATEVWLNKSMELL VSAGGLFOVALIRVLTGFEKKTGVYMQQ YPCYVDDIFLRVMSSSSPLFIMTHA AWTSYL IKGIVYEKEARLKETMRIMGLDNSLLWFSV SSLIPLLVSAGGLVVILKLGNLLPYSDPSVV FAYLQDVVEGALIRVLTGFEKKTGVYMQQ YPCYVDDIFLRVMSSSMELPTHALWYSRIKMGL VERTNKINGGYWDPOPRADPFEDMRTVINGS SSLIPLLVSAGGLVVILKLGNLLPYSDPSVV FLSVFAVVTILQCFLISTLFSRANLAAACGG YFITLYPPVLCVAWQDYVGFTKIKAGSLLS VARGGCEYFALFEEQGIGVQWDNLFERS PGOYGIPRPWYFPCTKSYWFGEESDEEKH PGOYGIPRPWYFPCTKSYWFGEESDEEKH NCRISEICMEEPTHLALLGYGMALKYSI NLGVCPQHNVLFDMLTVEEHITYPARLIG EKHVKABMEQMALDVGLPSSKLLKSTNG NLGVCPQHNVLFDMLTVEEHITYPARLIG GEKHVKABMEQMALDVGLPSSKLLKSTNG GRANLSVALAFVGGSKVULDEFTAGU GRANLSTULLKYQGGTIILSTHHINGGTGYD GDRAILSISHGLCCVGSSLFLKHOLIGTGYUL GRANLSTELGECCCYGSSLFLKHOLIGTGYUL GRANLSTELGECCCCCSSLFLKHOLIGTGYUL GRANLSTELGECCCCSSLFLKHOLIGTGYUL GRANLSTELGECCCCCCSSLFLKHOLIGTGYUL GRANLSTELGE	975	2325	A	8249		1571	LVALKNWKPKGTNIPAPOSPVEGEAVSGVVM
EEKGKYLPSLEMFRQRFRGGYHDTFGFR LSQLRVLCCEWLRPEHTKELBLLIVEQ LLPQELQAWVQEHCPESAEEAVTLLEDI DEPCHQWSTPPPEQKPWEKISSGTAKE SMQPQPLETSHKYESWGPLYIQESGEGEI DPRKVRDCRLSTQHESADEQKOSRAEGI DISVIANKPASLERQCVNLENEKGTKP EAGSKKGRESVPTKYTPGERRYICAEGGK, NSSNLTKHRRTHTGERPYVCCGGKAFGKOSLIR- HTGERPYQCKDGGKAFGKOSLIR- HTGERPYQCKDGGKAFGKOSLIR- HTGERPYQCKDGKAFGKOSLIR- HTGERPYQCKDGKAFGKOSLIR- HTGERPYQCKDGKAFSGKOSLIR- HTGERPYQCKDGKAFSGKOSLIR- HTGERPYQCKDGKAFSGKOSLIR- HTGERPYQCKDGKAFSGKOSLIR- HTGERPYQCKDGKAFSGKOSLIR- HTGERPYQCKDGKAFSGKOSLIR- HTGERPYQCKDGKAFSGKOSLIR- HTGERPYQCKDGKAFSGKOSLIR- HTGERPYQCKDGKAFSISHARIR- PWCHHCGKTFCSKSNLSKHQRVHTGEGI PWCHCHGKTFCSKSNLSKHQRVHTGEGI PWCHCHGKTFCKSSNLSKHQRVHTGEGI PWCHCHGKFFCKSSNLSKHQRVHTGEGI PWCHCHGKTFCKSSNLSKHQRVHTGEGI PWCHCHGKTFCKSSNLSKHQRVHTGEGI PWCHCHGKFFCKSNLSKHQRVHTGEGI PWCHCHGKTFCKSSNLSKHQRVHTGEGI PWCHCHGKTFCKSSNLSKHQRVHTGEGI FYWCHCHGKTFCKSSNLSKHQRVHTGEGI FYWCHCHGKTFCKSSNLSKHQRVHTGEGI FYWCHCHGKTFCKSSNLSKHQRVHTGEGI FYHLSIPRICHERSPTPTGEAP VGNRNSSIVARLFSDARRLLYSQKDTSNLSH KURLLYSQKDTILSTH KURLLYSQKDT KURLTSTATAT KURLTSTATAT KURLTSTATAT	1 1		ł	}		1971	MTKVLGMAPVLGPRPPOFOVGPLMVKVEEK
I ISQURVLCCEWLRFEHTKEQUELULVEQ I ILPQELQAWQEHCPESABEAYTILEDLEI DEPGIQVSTPPNEQKPWWEKISSSGTAKES SMQPQPLETSHYESWGPLYQESGEGEI DEPRKYRDCRLSTQHEESABEQKGSEAGEI DISVIIANNFEASLERQCVNLENEKGTKPP EAGSKKGRESVPYKTPTGEKRYLCAEGGK, NSSNLTKHRRTHTGERPYVCTKCGAFSH NLTLHYRFHLUDRPYDCKCGKAFGGKSDEAGH GEKPYKCKECGKAFNSKNSKHRHRHTGE PYWCHHCOKTFCSKSNLSKHQRVHTGEGI PYWCHHCOKTFCSKSNLSKHQRVHTGEGI PYWCHHCOKTFCSKSNLSKHQRVHTGEGI PYWCHHCOKTFCSKSNLSKHQRVHTGEGI PYWCHHCOKTFCSKSNLSKHQRVHTGEGI PYWCHHCOKTFCSKSNLSKHQRVHTGEGI PYWCHHCOKTFCSKSNLSKHQRVHTGEGI PYWCHHCISVRLSYPYEQHECHFPNK PSAGTLPWQGIGNANNPCFRYPTTGEAP VONFNKSTVALFSDARFLLYSQKDTSMI MRKVLRTLQQKSSSSNLKLQDFLVDNETI GFLYHNLSLPKSTVDKMLRADVLHKYFL YQLHLTSLCNGKSSESMLKLQDFLVDNETI GFLYHNLSLPKSTVDKMLRADVLHKYFL YQLHLTSLCNGKSSESMLKLQDFLVDNETI GFLYHNLSLPKSTVDKMLRADVLHKYFL YQLHLTSLCNGKSSESMLKLQDFLVDNETI GFLYHNLSLPKSTVDKMLRADVLHKYFL YQLHLTSLCNGKSSESMLKLQDFLVDNETI GFLYHNLSLPKSTVDKMLRADVLHKYFL YQLHLTSLCNGKSSESMLKLQDFLVDNETI GFLYHNLSLPKSTVDKMLRADVLHKYFL YQLHLTSLCNGKSSESMLKLQDFLVDNETI GFLYHNLSLPKSTVDKMLRADVLHKYFL YQLHTSLCNGKSSESMLKLQDFLVDNETI GFLYHNLSLPKSTVDKMLRADVLHKYFL YQLHTSLCNGKSESMLKLQDFLVDNETI GFLYHNLSLPKSTVDKMLRADVLHKYFL YQLHTSLCNGKSSSNLKLQDFLYBKSSSNLKLQDFLYDNETI GFLYHNLSLPKSTVDKMLRADVLHKYFL YQLHTSLCNGKSSSSTURA PSTAMENTAL PSTAMENTAL REMARKATION REM]				
ILPQELQAWVQEHCPESAEAEATLLEDLE DEPGHQVSTPPNEQKP VWEKISSSGTAKE SMOPQPLETSHKYESWGPLYIQESGEEGEI DPRKVRDCRLSTQHEESADEQKGSEAGGL DIBSVIIANKPEASLERQCVNLENEKGTKEP EAGSKKGRESVPTKPTPGERRYICAECGK NSSNLTKHRTHTGERFYVCKDCGKAFGQSSDI HORMHTEEAPYQCKDCGKAFGGSSDI HORMHTEEAPYQCKDCGKAFGGSSDI HORMHTEEAPYQCKDCGKAFGGSSDI HORMHTEEAPYQCKDCGKAFGGSSDI HORMHTEEAPYQCKDCGKAFGGSGSLI HITGEKPYQCNECGKSFSQHAGLSSHQRLI GEKPYKCKECGKAFNHSSNFNKHRIHTG PYWCHHCGKTFCKSNSINSKHQRYHTGEGI PYWCHHCGKTFCKSNSINSKHQRYHTGEGI PYWCHHCGKTFCKSNSINSKHQRYHTGEGI GEKPYKCKECGKAFNHSSNFNKHRIHTG PYWCHLCGKTFCKSNSINSKHQRYHTGEGI GEKPYKCKECGKAFNHSSNFNKHRIHTG PYWCHLCGKTFCKSNSINSKHQRYHTGEGI GFKPYKCKECGKAFNHSSNFNKHRIHTG PYWCHLCGKTFCKSNSINSKHQRYHTGEGI GFKPYKCKECGKAFNHSSNFNKHRIHTG GFKPYKKLCCGKAFNHSSNFNKHRIHTG GFKPYKCKECGKAFNHSSNFNKHRIHTG GFKPYKCKECGKAFNHSSNFNKHRIHTG GFKPYKCKECGKAFNHSSNFNKHRIHTG GFKPYKCKECGKAFNHSSNFNKHRIHTG GFKPYKCKECGKAFNHSSNFNKHRIHTG GFKPYKCKECGKAFNHSSNFNKHRIHTG GFKPYKCKECGKAFNHSSNFNKHRITG GFKPYKCKECGKAFNHSSNFNKHRITG GFKPYKCECGKAFNHSSNFNKHRITG GFKPYKCKECGKAFNHSSNFNKHRITG GFKPYKCKECGKAFNHSSNFNKHRITG GFKPYKCKECGKAFNHSSNFNKHRITG GFKPYKCKECGKAFNHSSNFNKHRITG GFKPYKCKECGKAFNHSSNFNKHRITG GFKPYKCKECGKAFNHSSNFNKHRITG GFKPYKCKECGKAFNHSSNFNKHRITG GFKPYKCKECGKAFNHSSNFNKHRHTG GFKPYKCKE	i						
DEPGHQVSTPPNEOKPYWEKISSGTAKES SMOPOPLETSHKYESWOT/JOEGSGEOEG DPRKVRDCRLSTQHEESADEQKGSEAEGL DISVIIANKPEASLERQCVNLENEKGTKPP EAGSKKGRESVPYKTPGEKGKAFSH NLTLHYRHLVDRPYDCKCGKAFGGKSDI.RH INTGEKPYQCKCGKAFGGKSDI.RH IHTGEKPYQCCKDCGKAFSGKGSL.RH IHTGEKPYQCCMCGKSFSQHAGI.SSNORLH GEKPYKCKEGGKAFNSNNNKHRHRHTG PYWCHHCGKTFCSKSNLSKHQRVHTGEGI P 976 2326 A 8257 298 7086 GMAACWPQLRLLLWKNLTFRRRQTCQLL VAWPLFIELLISWILSYPQHECHPPNK PSAGTLPWVQGIICNANNPCFRYPTIGGE VONFNKSIVARLFSDARRLLLYSQKDTSM MKRVLRTLQOKKSSSLAULOPFLVDNETI GFLYHNLSJEKSTYDKMLRADVILHKVFL YQLHLTSLCNGKSSEEMLQDFLVDNETI GFLYHNLSJEKSTYDKMLRADVILHKVFL YQLHLTSLCNGKSSEEMLGDFLVDNETI GFLYHNLSJEKSTYDKMLRADVILHKVFL YQLHLTSLCNGKSSEEMLGDFLVDNETI GFLYHNLSJEKSTYDKMLRADVILHKVFL YQLHLTSLCNGKSSEEMLGDFLVDNETI GFLYHNLSJEKSTYDKMLRADVILHKVFL YQLHLTSLCNGKSSEEMLGDFLVDNETI GFLYHNLSJEKSTYDKMLRADVILHKVFL YQLHLTSLCNGKSSEEMLGDFLVDNETI GFLYHNLSJEKSTYDKMLRADVILHKVFL YQLHTSLCNGKSSEEMLDDFLVDNETI GFLYHNLSPKSTYDKMLRADVILHKVFL YQLHTSLCNGKSSEEMLDDFLVDNETI GFLYHNLSPKSTYDKMLRADVILHKVFL YQLHTSLCNGKSSEEMLDDFLURT KVAPPLFURDNETI GFLYHNLSSSSSTIPHTWAN PLAVGKILYTPDTPATRQVMAEVNKTFQEI VFPHDLEGGMWELSFKLINKSNECLDD MLDSRDNDHFWEQQLDGLDWTAQDTVA AKHPDVQSNGSYYTWREAFISTRQNAR PLAVGKILYTPDTPATRQVMAEVNKTFQEI VFPHDLEGGMWELSFKLINKSNECLDD KFWAGIVFTGITPGSIELPHHVKYKIRNGID VFRTNKIKDGYWDFGPFDMRYVW FAYLQDVVEQAIRVLTGTEKKTQVYMQQ YPCYVDDIFLRVMSSMLIPHSNSWSDSVYTWREAFISTRQAIR RFPMECVNLNKLEPIATEVNSMELDSIBLNSWS SSLIPLLVSAGLLVVILKLGNILDFYSDSVV FAYLQDVVEQAIRVLTGTEKKTQVYMQQ YPCYVDDIFLRVMSSMLFDTFLAVWYSV. IKIGVYFKEARLKETMRIGLDNSIL WFSV SSLIPLLVSAGLLVVILKLGNILDFYSDSVV FAYLQDVVEQAIRVLTGTEKKTQVMINYSV. IKIGVYFKEARLKETMRIGLDNSIL WFSV SSLIPLLVSAGLLVVILKLGNILDFYSDSVV FLSVFAVVTILQCFLISTRANLAAACGG YFTLYIPYVLCVAWQDVVGFTLKIFRSILS VAFGFGCEYFALFEEOGGRESHFHKORGK TTMSILTGLFPFTGTATVILGKDRESBSTRR NLGVCPQHNVLFDMLTVEEHWFYARLGG GGMQRKLSVALAFVGGSKVVILDETFAGV YSRRGIWELLLKYRQGRTILLSTHIMDEAD GGMARKLSVALAFVGGSKVVILDETFAGV YSRRGIWELLLKYRQGRTILLSTHIMDEAD GGMARKLSVALAFVGGSKVVILDETFAGV YSRRGIWELLLKYRQGRTILLSTHIMDEAD GGMARKLSVALAFVGGSKVVILDETFAGV	1 1						ILPOELOAWVOEHCPESAEFAVTI LEDI EREL
SMOPOPLETSHKYESWOPLYIOESGEEGEI DPRKVRDCRLSTQHEESADEQKGSEAGEI DISVIIANKPEASLERQCVNLENEKGTKPP EAGSKKGRESVPTKPTPGERRYICAEGCS/ NSSNLTKHRRTHTGEKPYVCTK.CGKAFSH NLTLHYRTHLVDRRYDCK.CGKAFGQSSDI HORMHTERAPYQCKDCGKSFSGKGSLIKH GEKPYKCKECGKAPTISTRYGKGSTSH HTGEKPYQCNEGGKSFSQHAGLSSHORLI GEKPYKCKECGKAPTISTRYHEAPT PWCHHCGKTFCSKSNLSKHQRVHTGEGI P 976 2326 A 8257 298 7086 GNMACWPQLRLLLWKNLTFRRRQTCQLL VAWPLFIFLILISVRLSYPPYEQHECHPNK PSAGTLPWYOGIICNANPNKHRHTHTO PYWCHHCGKTFCSKSNLSKHQRVHTGEGI P 976 P 1086 GNMACWPQLRLLLWKNLTFRRRQTCQLL VAWPLFIFLLISVRLSYPPYEQHECHPNK PSAGTLPWYOGIICNANPKLHENTHTO PYWCHHCGKTFCSKSNLSKHQRVHTGEGI P 1087 GRANTLLSCHSTARTLLLYSQKDTSM MRKVLRTLQUKKSSSNLKLQDFLVDNETI GFLYHNLSLPKSTVDKMLRADVILHKVFLI YQLHLTSLCNGSKSEEMIQLGDGEVSELCC REKLAAARRVLRSNMDLHRTLNTSTSPI KELABATKTLLHSLGTLAQELFSMRSWSD QEVMFLTNVNSSSSTQLYASRIVCAP FEKLAAARRVLRSNMDLHSTRINTLNSTSPI KELABATKTLLHSLGTLAQELFSMRSWSD QEVMFLTNVNSSSSTQLYASRIVCAP FELAVARVLTSNMDLFDTATRQVMAEVNKTFQEI GGLKISLNWYEDNNYSALFGGNOTEED TTYDNSTTPVCNDLMKNLESPLSTRINKA PLLWGKILYTPDTATRQVMAEVNKTFQEI FENDANDHFWEQQLUSDLDWTADDIVA AKHPEDVQSSNGSVYTWREAFINETNQAR REPMECVNLINKLEPIATEVRANTENSSMELD KFWAGIVFTGITPGSIELPHHVKYKIRMGIE VERTNIKLGDYWDGPRADFEDMRYVW FAYLQDVVEQAIRVLITEKKTICVYMQQ YPCYVDDIFLRYMSRSMPLFMTLAWYSV. IIKGVYYSKEARLKETNKINGLDNSIL WFSV SSLIPLLVSAGLLVVILKLGNLLPYSDPSVV FLSYFAVVTLQCFLISTTEKKTGVYMQQ YPCYVDDIFLRYMSRSMPLFMTLAWYSV. IIKGVYSKEARALKETNKINGLDNSIL WFSV SSLIPLLVSAGLLVVILKLGNLLPYSDPSVV FLSYFAVVTLQCFLISTFARALAAACGG YFTLYLPYVLCVAWQDVVGFTLKIFASLLS VAFGGCEYFALFEEQGIGVQWDMLFISPV EDGFNLTTSVSMMLFDTLYGMTWYIEA PGQVGIPRPWYFPCTKSYWFGEESBCSHF NGKVADGLALNYFEGOTISLGHNGAGK TTMSLTGLEFPTSGT AYLICKDIRSBMSTTR NLGVCPQHNVLFDMLTVEEHWFYARLKG EKHVKAEMEQMALDVGLPSSKLKSKTSQL GGMQRKLSVALAFVGGSKVVILDETAGV YSRRGIWELLLKYRQGRTILLSTHIMDEAD GDRAINSHOKLCCVGSSLYLKNOLGTGYYY GRKVALDELLLKYRQGRTILLSTHIMDEAD GDRAINSHOKKLCCVGSSLYLKNOLGTGYYY GRKVALDELLLKYRQGRTILLSTHIMDEAD GDRAINSHOKKLCCVGSSLYLKNOLGTGYYY	1 1		ľ	1			DEPGHOVSTPPNEOKPVWEKISSSGTAKESPS
DPRKVRDCRLSTOHEESADEQGSEAEGI DISVILANKPEASLERQCYNLENEKGTKPP EAGSKKGRESVPTKPTPGERRYICAECGK, NSSNLTKHRRTHTGEKPYVCTKCGKAFSH NLTLHYRTHLVDRPYDCKCGKAFGGSSDI HQRMHTEEAPYQCKCDCKAFGGGSLIKH HITGEKPYVCTKCGKAFSH HITGEKPYVCTKCGKAFSH HITGEKPYVCTKCGKAFSH HITGEKPYVCTKCGKAFSGSLIKH HITGEKPYVCTKCGKAFSH HITGEKPYVCTKCGKAFSH GERPYKCKEGGKAFNHSSNFNKHRHTHTG PYWCHHCGKTFCSKSNLSKHQRHTGGIG GERPYKCKEGGKAFNHSSNFNKHHRHTHTG PYWCHHCGKTFCSKSNLSKHQRHTGHTPIK GERPYKCKEGGKAFNHSSNFNKHHRHTHTG PYWCHHCGKTFCSKSNLSKLQDFLVDKET GERPYKCKEGGKAFNHSSNFNKHRHTHTG PYWCHHCGKTFCSKSNLSKLQDFLVDNET GFLYHNLSLPRSTYDKMLRADVILHKVFLL YQLHLTSLCNGSKSESSNLKLQDFLVDNET GFLYHNLSLPRSTYDKMLRADVILHKVFLL YQLHLTSLCNGSKSESSNLKLQDFLVDNET GFLYHNLSLPRSTYDKMLRADVILHKVFLL YQLHLTSLCNGSKSSSNLKLQDFLVDNET GFLYHNLSLPRSTYDKMLRADVILHKVFLL YQLHLTSLCNGSKSSNLKLQDFLVDNET GFLYHNLSLPRSTYDKMLRADVILHKVFL YQLHLTSLCNGSKSSNLKLQDFLVDNET GFLYHNLSLPRSTYDKMLRADVILHKVFL YQLHLTSLCNGSKSSNLKLQDFLVDNET GFLYHNLSLPRSTYDKMLRADVILHKVFL YQLHLTSLCNGSKSSNLKLQDFLVDNET GFLYHNLSLPRSTYDKMLRADVILHKVFL YQLHLTSLCNGSKSSNLKLQDFLVDNET GFLYHNLSLPRSTYDKMLRADVILHKVFL YQLHLTSLCNGSKSSNLKLQDFLVDNET GFLYHNLSLPRSTYDKMLRADVILHKVFL GGKKLSLNWYEDNNYKLTFGGI VFHDLEGMWEELSPKIWFMENSORDLD KEWAGIVFTGTTPGSIELPHHVYKRIMGIL VERTNIKIRGDYWDPGPRADPFEDMRYVW FAYLQDVUSQAIRVLTGTEKKTQVYMQQ YPCYVDDIFLRVMSSSMSTLFWINSMELLDI KEWAGIVFTGTTGSELPHHVYKRIMGIL VERTNIKIRGDYWDPGPRADPFEDMRYVW FAYLQDVUSQAIRVLTGTEKKTQVYMQQ YPCYVDDIFLRVMSSSMSTLFMINGLNSLUKFSSPL URGVYEKEARL KETMRIMGLDNSL WFSV SSLIPLLVSAGLLVVILKLGHLIPLSDFSVL VAGFGCCFYALFEGGGYOWDNLFESP EDGFNLTTSVSMMLFDTFLYGVMTWYIEA PQQYGIPRWYPCTKSSYWFGEBDEKSHI NIGVCPQHNVLFDMLTVEEHHWYARLG EKHVKAEMEQMALDVGLIPSSKLKSKTSQL GGMQRKLSVALAFVGGGSVVLIDEFTAGW YSRRGIWELLLKYRQGRTILSTHHMDEAD GDRAIISHGKLCCVGSSLFLKNOLGTGYYI GGMQRKLSVALAFVGGSKVVLIDEFTAGW YSRRGIWELLLKYRQGRTILSTHHMDEAD GDRAIISHGKLCCVGSSLFLKNOLGTGYYI GGMARGLSVALAFVGGSRTILSTHHMDEAD GDRAIISHGKLCCVGSSLFLKNOLGTGYYI							SMOPOPLETSHKYESWGPLYTOESGEEOFFAC
DIISVIIANKPEASLERQCVNLEKGTKYP EAGSKGRESVPTKPFOERRYICAECGKA NSSNLTKHRRTHTGEKPYVCTKCGKAFSH NLTLHYRTHLVDRPYDCKCGKAFGQSSDLRH HQRMHTEEAPYQCKDCGKAFSGKGSLIRH HITGEKPYQCNECGKSFSQHAGLSSHORLH GERPYKCKEGKAFNSHNFNHRHHTG PYWCHHCGKTFCSKSNLSKHQRVHTGEGI PYWCHHCGKTFCSKSNLSKHQRVHTGEGI PYWCHHCGKTFCSKSNLSKHQRVHTGEGI PYWCHILLWKNLTFRRRQTCQLL VAWPLFIELILSVRLSYPPYGHEGLFPPNK PSGATLPWVQGIGNLHLWKNLTFRRRQTCQLL VAWPLFIELILSVRLSYPPYGHEGLFPNK PSGATLPWVQGIGNLMRADVILHKVFL YQLHLTSLCNGSKSEEMIQLGDQEVSELCC REKLAAAERVLRSNMLRADVILHKVFL YQLHLTSLCNGSKSEEMIQLGDQEVSELCC REKLAAAERVLRSNMLRADVILHKVFL YQLHLTSLCNGSKSEEMIQLGDQEVSELCC REKLAAAERVLRSNMLRADVILHKVFL YQLHLTSLCNGSKSEEMIQLGDQEVSELCC REKLAAAERVLRSNNDNYRALFGGNGTEED GGKJKISLNWYEDNNYKALFGGNGTEED TFYDNSTTPYCNDLMKNLESSPLSRIIWKA PLLVGKILYTPTDFATVQMAEVNKTFQEI GGKJKISLNWYEDNNYKALFGGNGTEED TFYDNSTTPYCNDLMKNLESSPLSRIIWKA PLLVGKILYTPTDFATVQMAEVNKTFQEI FYFDNSTTPYCNDLMKNLESSPLSRIIWKA PLLVGKILYTPTDFATVAMAEVNKTFQEI FYFDNSTTPYCNDLMKNLESSPLSRIIWKA PLLVGKILYTPTDFATVAMAEVNKTFQEI FYFDNSTTPYCNDLMKNLESSPLSRIIWKA PLLVGKILYTPTDFATVAMAEVNKTFQEI FYFDNSTTPYCNDLMKNLESSPLSRIIWKA PLLVGKILYTPTDFATVAMAEVNKTFQEI FYFDNSTTPYCNDLMKNLESSPLSRIIWKA PLLVGKILYTPTGATATEVNAMETEN FYFDNSTTPYCNDLMKNLESSPLSRIIWKA PLLVGKILYTPTGATATEVNAMETEN FYFDNSTTPYCNDLMKNLESSPLSRIIWKA PLLVGKILYTPTGATATEVNAMETEN FYFDNSTTPYCNDLMKNLESSPLSRIIWKA PLLVGKILYTPTGATATEVNAMETEN FYFDNSTTPYCNDLMKNLESSPLSRIIWKA PLLVGKILYTPTGATATEVNAMETEN FYFDNSTTPYCNDLMKTTPTT FYNNSTTPYCNDLMKNLESSPLSTIIWKA PLLVGKILTSVSSMLLGDTLTNNSMELLDI KFWAGIVFTGTTOSTT AKHTEN FYFDNSTTPYCNCH FY	1						DPRKVRDCRLSTOHEESADEOKGSEAEGLKG
EAGSKKGRESVPTKPTPGERRYICAECGKA NSSNLTKHRRTHTGERYPVCTKCGKAFSH NLTLHYRTHLVDRPYDCKCGKAFGQSDI HQRMHTBELAPYQCKDCGKAFSGKGSLIRH HTGEKPYQCNECGKSFSQHAGLSSHQRLI GERPYKCKEGGKAFNHSSNFNKHRHHRHTG PYWCHHCGKTFCSKSNLSKHQRHTGGGI PYWCHHCGKTFCSKSNLSKHQRHTGGGI PYWCHHCGKTFCSKSNLSKHQRHTGGGI PYWCHICGKTFCSKSNLSKHQRHTGGGI PYWCHICGKTFCSKSNLSKHQRHTGGGI PYWCHICGKTFCSKSNLSKHQRHTGGGI PYWCHICGKTFCSKSNLSKHQRHTGGGI PYWCHICGKTFCSKSNLSKHQDFLVDTGGI GRAMACWPQLRLLLWKNLTTFRRQTCQLL VAWPLFFLLLISVBLSYPPYEQHECHFPNK PSAGTLPWVQGIICNANNPCFRYTFGEAP VGNFNKSIVARLFSDARRLLLYSQKDTSMI MRKVLRTLQQIKKSSSNLKLQDFLVDNETI GFLYHNLSLPKSTVDKMLADLHSVHSTVOKI YQLHLTSLCNGSKSEEMIQLGDQEVSELCC REKLAAAERVLRSNMDILKPILTILNSTSPJ KELAEATKTLHSLGALFSMRSWSD QEVMFLTNVNSSSSTQIVQAVSRIVCGHP GGLKIKSLNWYEDNNYKALFGGNGTEED/ GGLKIKSLNWYEDNNYKALFGGNGTEED/ TYPDNSTTPYCNDLMKNLESSPLSRIWKA PLLVGKILYTPDTPATRQVMACHVKTTQGI VFHDLEGMWEELSPKIWTFMENSQEMDLU MLDSRDNDHFWEQQLDGLDUNTAQDIVA AKIPEDVQSSNGGSVYRALFGGNFTEDMRYVW FAYLQDVVEQAIRVLTGTEKKTQYYMQQ YPCYVDDJFLRVMSKRMELLDI KEWAGIVFTGITPGSIELPHHVKYRIMGIL VERTNKIKGDGYWDPGPRADPFEDMRYVW FAYLQDVVEQAIRVLTGTEKKTQYYMQQ YPCYVDDJFLRVMSRMPLMTLAWIYSV, IKGIVYEKEARLKETMRIMGLDNSILWFSN SSLIPLLVSAGLLVVILKLGNLLYFSDPSVV FLSVFRAVVTILQCFLISTLFSRANLAAACCG YFTLYLPYVLCVAWQDYVGFTLKIFASLLS VAAFGGCEYFALFEGGGGTVLKIFASLLS VAAFGGCEYFALFEGGGGGTVKLIFASLLS VAAFGGCEYFALFEGGGGGVQWDNLFESP EGGFNLTTSVSSMMLPDTFLYGWMTWYIEA PQQYGJPRPWYFPCTKSYWFGGESDEKSHP NQKRISEICMEEPTHLKLGVSIQNLVKVY GMKVAVDGLALNFYEQGITSTLKIFASLLS TTMSLTGJPPPTSGTAYLIGKDRSEMSTIR NLGVCPQHNVLFDMLTVEEHHWFVARLKG EKHVKAEMEQMALDVGLPSSKLKKNTSQL GGMQRKLSVALAFVGGGSVVILDEFTAGV YSRRGIWELLLKYRQGRTILLSTHHMDEAD GDRAIISHGKLCCVGSSLFLKNOLGTGYYI	j j			i			DIISVIIANKPEASLEROCVNI ENEKGTKPPI O
NSSNLTKHRRTHTGEKPYVCTKCGKAFGQSSDI HULTHYRTHLVDRYPDCKCGKAFGQSSDI.HULTHYRTHLVDRYPDCKCGKAFGQSSDI.HULTHYRTHLVDRYPDCKCGKAFGGSSLI.HULTGEKPYQCNECGKSFSGKGSLI.HULTGEKPYQCNECGKSFSGKAGL.SKH.HIRTIG GEKPYCKECGKAFNHSSNFNKHHRIHTG PYWCHHCGKTFCSKSNLSKHQRVHTGEGI PYWCHHCGKTFCSKSNLSKHQRVHTGEGI PYWCHHCGKTFCSKSNLSKHQRVHTGEGI PYWCHHCGKTFCSKSNLSKHQRVHTGEGI PYWCHHCGKTFCSKSNLSKHQRVHTGEGI PYWCHHCGKTFCSKSNLSKHQRVHTGEGI PYWCHHCGKTFCSKSNLSKHQRVHTGEGI PYWCHHCGKTFCSKSNLSKHQRVHTGEGI PYWCHLGKTSVRJENGLADELULYSQKDTSMI MRKVLRTLQUIKKSSSNLKLQDFLVDRIEG GFLYHNLSLPKSTVDKMLRADVILHKVFLL YQLHLTSLCNGSKSEEMIQLGDQEVSELCC REKLAAAFRVLRSSMDILKPILRTINSTSPI KELAAAFRVLRSSMDILKPILRTINSTSPI KELAAAFRVLRSSMDILKPILRTINSTSPI KELAAAFRVLRSSMDILKPILRTINSTSPI KELAAAFRVLRSSMDILKPILRTINSTSPI KELAAAFRVLRSSMSDILKPILRTINSTSPI KELAAAFRVLRSSMSDILKPILRTINSTSPI KELAAAFRVLRSSMSDILKPILRTINSTSPI KELAAAFRVLRSSMSDILKPILRTINSTSPI KELAAAFRVLRSSMSDILKPILRTINSTSPI KELAAAFRVLRSSMSDILKPILRTINSTSPI KELAAAFRVLRSSMSDILKPILRTINSTSPI KELAAAFRVLRSSMSDILKPILRTINSTSPI KELAAAFRVLRSSMSDILKPILRTINSTSPI KELAAAFRVLRSSMSDILKTINSTSPI KELAAAFRVLRSSMSDILKTINSTSPI KELAAAFRVLRSSMSDILKTINSTSPI KELAAAFRVLRSSMSDILKTINSTSPI KELAAAFRVLRSSMSTITMALTINSTVOMARTINSTVILTINSTSPI KELAAAFRVLRSSMSTITMALTINSTVILTINSTPI KELAAAFRVLRSSMSTITMALTINSTVILTINSTPI KELAAAFRVLRSSMSTITMALTINSTVILTINSTPI KELAAAFRVLRSSMSTITMALTINSTVILTINSTSPI KELAAAFRVLRSSMSTITMALTINSTVILTINSTSPI KELAAAFRVLRSSMSTITMALTINSTVILTINSTSPI KELAAAFRVLRSSMSTITMALTINSTVILTINSTPI KELAAAFRVLRSSMSTITMALTINSTSVSMMLFDTITLAWTSV. IKGIVYEKAALFVGGIVQWDNILFISP EDGFNLTTSVSMMLFDTITLAWTSV. IKGIVYEKAALFVGGIVQWDNILFISP EDGFNLTTSVSMMLFDTITLAWTSV. IKGIVYEKAALFVGGIVQWDNILFISP EDGFNLTTSVSMMLFDTITLAWTSV. IKGIVYEKAALFVGGIVQWDNILFISP EDGFNLTTSVSMMLFDTITLAWTSV. IKGIVYEKAAMFQDALINGTSCHITLSTHHMDEAD GGGMGRKLSVALAFVGGSKVULDEPTAGV. YSRRGIWELLLKYRQGRTILLSTHHMDEAD GGRANLISHGKLCCGSSLFLKNOLGTGYYI							
NLTLHYRTHLVDRPYDCKCGKAFGQSSDLIRHQRMHTEEAPYQCKDCGKAFSGKGSLIRHGERYYQCKDCGKSFSQHAGLSSHQRLIGERYYQCKDCGKSFSQHAGLSSHQRLIGERYYCKCECGKSFNHSSNFNKHHRIHTGERPYWCHCGKFFCKKSFSCHAGLSSHQRLIGERYYKCKECGKAFNHSSNFNKHHRIHTGERPYWCHCGKFFCKSKSNLSKHQRVHTGGGIPWCHCHFFNKKSTVARLFSGAGTLPWVQGICNANNPCFRYPTGEAPVGNFNKSIVARLFSDARRLLLYSQKDTSMMRKVLRTLQUKKSSSNLKLQDFLVDNETIGFLYNKSIVARLFSDARRLLLYSQKDTSMMRKVLRTLQUKKSSSNLKLQDFLVDNETIGFLYNKSIVARLFSDARRLLLYSQKDTSMMRKVLRTLQHSKSSSNLKLQDFLVDNETIGFLYNTSPHKSIVARLFSGNGGLGDCVSELCCAREKLAAAERVLRSNMDILKPILRTLNSTSPHKELAAAERVLRSNMDILKPILRTLNSTSPHKELAAAERVLRSNMDILKPILRTLNSTSPHKELAAAERVLRSNMDILKPILRTLNSTSPHKELAAAERVLRSNMDILKPILRTLNSTSPHKELAAAERVLRSNMDILKPILRTLNSTSPHKELAAAERVLRSNMDILKPILRTLNSTSPHKELAAAERVLRSNMDILKPILRTLNSTSPHKELAAAERVLRSNMDILKPILRTLNSTSPHKELAAAERVLRSNMDILKPILRTLNSTSPHKELAAAERVLRSNMDILKPILRTLNSTSPHKELAAAERVLRSNMDILKPILRTLNSTSPHKELAAAERVLRSNMDILKPILRTLNSTSPHKELAAAERVLRSNMDILKPILRTLNSTSPHKELAAAERVLRSNMVEDDARVAWAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA						•	
HQRMHTEEAPYQCKDEGKAFSGKAFSGKGSLIZHE HTIGERPYQCNECKKSFSQHAGLSSHQRLI GEKPYKCKECGKAFNHSSNFNKHRRIHTG PYWCHHCGKTFCSKSNLSKHQRVHTGEGI PYWCHHCGKTFCSKSNLSKHQRVHTGEGI P GNMACWPQLRLLLWKNLTFRRRQTCQLL VAWPLFIFLILISVRLSYPPYEQHECHFPNK PSAGTLPWQGIICNANNPCFRYPTPGEAP VGNFNKSIVARLFSDARRLLLYSQKDTSMI MRKVLRTLQQIKKSSLEKIQIGDQEVSELIC EFLYANLSIPKSTYDKMLRADVILHKVFL YQLHLTSLCNGSKSEEMIQLGDQEVSELIC REKLAAAERVLRSSMDILKPIRTILNSTSPI KELAEATKTLHSLGTLAGELFSMRSWSDI QEVMFLTNVNSSSSTQIVQAVSRITVCGHP GGLKIKSLNWYEDNNYKALFGGNGTEED/ TFYDNSTTPYCNDLMKNLESPLSRIIWKA PILVGKILYTPDTPATRQVMAEVNKTFQEI VFHDLEGMWEELSPKIWTFMENSQEMDIJ MLDSRNDHFPWGQLDGIDUFAVRITVGHP VFHDLEGMWEELSPKIWTFMENSQEMDIJ AKHPEDVQSSNGSVYTWREAFNETNQAIR RFMECVVLNKLEPIATEVWLNKSMELLDI KFWAGIVFTGITPGSIELPHHVKYKRMGIE VERTNKIKDGYWDPGPRADPFEDMRYVW FAYLQDVVEQAIRVLTGTEKKTGVYMQQ YPCYVDDIFLRVMSSSMPLFMTLAWIYSV. IIKGIVYEKEARLKETMRINGLDNSILWFSS SLIPPLLVSAGLLVVILKGNLLPYSDPSVV FLSVFAVVTILQCFLISTLFSRANLAAACGG YFTLYLPYVLCVAWQDYVGFTLKIFASLLS VARGFGCEYFALFEEQGIGVQWDNLFESPV EDGFNLTTSVSMMLFDTFLYGVMTWYIEA PGQVGPRPWYFPCTKSYWFGEESDEKSHP NQKRISEICMEEEPTHLKLGVSIQNLVKVYI GMKVAVDGLALNFYEGGITSLGHNGAGK TIMSILTGLFPTSGTAYILGKDRESMSTIR NLGVCPOHNVLFDMLTVEEHWFYARLKG EKHVKAEMEQMALD VGLPSSKLK KSKTSOL GGMQRKLSVALAFVGGSKVVILDEPTAGV YSRRGIWELLLKYRQGRTIILSTHHMDEAD GDRIAIISHGKLCGGTGYYI							
IHTGEKPYQCNECGKSFSQHAGLSSHQRLI GEKPYKCKECKANHSSNFNKHRHRITG GYWCHCGKTCSKSNLSKHQRVHTGEG PYWCHCGKTCSKSNLSKHQRVHTGEG PYWCHCGKTCSKSNLSKHQRVHTGEG P	1 1						HORMHTEEAPYOCKDCGKAFSGKGSLIRHYR
976 2326 A 8257 298 7086 GNMACWPQLRLLLWKNLTFRRQTCQLL VAWPLFFLLISVRLSYPPYEQHECHFPNK PSAGTLPWVQGIICNANNPCFRYPTPGEAP VGNFNKSIVARLFSDARRLLLYSQKDTSM MRKVLRTLQOIKKSSSNLKLQDFLVDNETI GFLYHNLSLPKSTVDKMLRADVILHKVFLI YQLHLTSLCNGSKSEEMQLGDQEVSELCC REKLAAAERVLRSNMDLLSPILTLINSTSP) KELAEATKTILHSLGTLAQELFSMRSWSD QEVMFLTNVNSSSSSTQIYQAVSRIVCGHP GGLKKSLNWYEDNNYKALFGGNGTEEDD TFYDNSTTPYCNDLMKNLESSPLSRIWKA PLLVGKILYTPDTPATRQVMAEVNKTFQEI VFHDLEGMWELSPKLWTFMENSGEMDLL KFWAGIVFTGITPGSIELPHHVKYKRMGIE VERTNKIKDGYDFOPMOFPADPFEDMRYVW FAYLQDVVEQAIRVLTGTEKKTGVYMQQ YPCVVDDIFLRVMSRSMPLFMTLAWIYSV. IKIGVYEKAEALKETMRIMGLDNSILWFSS SSLIPLLVSAGLLVVILKLGNLLPYSDPSVV FLSVFAALFERQEGGVQWDNLTESPY EDGFNLTTSVSMMLFDTFLYGVMTWYIEA PGQVGIPRWYPCTKSYWFGEESDEKSHP NQKRISEICMEEDTHLKLGVSIQNLVKVYI GMKVAVDGLALNFYEGQTISFLGHNGAKG TTMSILTGLFPPTSGTAYILGKDRSEMSTIR NLGVCPQHNVLFDMLTVEEHWFYARLKG EKHVKAEMEGDALNFYEGGTISFLGHNGAKG TTMSILTGLFPPTSGTAYILGKDRSEMSTIR NLGVCPQHNVLFDMLTVEEHWFYARLKG EKHVKAEMEGDALDVFLESPY GGMQRKLSVALAFVGGSKVVILDEPTAGV YSRRGIWELLLKYPQGGTSKLVILGLGTGTYG							IHTGEKPYOCNECGK SESOHAGI SSHORI HT
976 2326 A 8257 298 7086 GNMACWPQLRLLLWKNLTFRRRQTCQLL VAWPLFIFLILISVRLSYPPYEQHECHFPIK PSAGTLPWVQGIICNANNPCFRYPTPGEAP VGNFNKSIVARLFSDARRLLLYSQKDTSM MRKVLRTLQQIKKSSSNLKLQDFLVDNETI GFLYHNLSLPKSTVDKMLRADVILHKVFLI YQLHLTSLCNGSKSEEMIQLGDQEVSELCC REKLAAAERVLRSNMDILKPILRTLNSTSPI KELAEATKTILHSLGTLAQELFSMRSWSD QEVMFLTNNVSSSSTQIYQAVSRIVCGHP GGLKIKSLNWYEDNYXALFGGNGTEEDA TFYDNSTTPYCNDLMKNLESSPLSRIIWKA PLLVGKILYTPDTPATRQVMAEVNKTFQEI VFHDLEGMWEELSPKJWTFMENSQEMDLN MLDSRDNDFHEWQQDLGDUTAQDIVA AKHPEDVQSSNGSVYTWREAFNETNQALR RFMECVNLNKLEPIATEVWLINKSMELDB KFWAGIVFTGITSELPHHVKYKIRMGIL VERTNKIKDGYWDPGPRADPFEDMRYVW FAYLQDVVEQAIIRVLITGTEKKTGYVMQQ YPCYYDDIFIRLWSSSSMPLFMTLAWISVS, IKGIVYEKEARLKETMRIMGLDNSILWFSV SSLIPLLVSAGLLVULKLGNLLPYSDPSVV FLSVFAVVTILQCLTSTLFSRANLAAACAG YFTLYLPYVLCVAWQDYVGFTLKIFASLLS VAFGFGCEYFALFEEQGIGVQWDNLFESPV EDGFNLTTSVSMMLFDTFLYGVMTWYIEA PGQYGIPRWYFPCTKSYWFGEESDEKSHP NQKRISEICMEEEPTHLKIGVSIQNLVKVYI GMKVAVDGLALNFYEGQTTSFLGHNGAGK TTMSILTGLFPPTSGTAYILGKDRSEMSTIR NLGVCPQHNVLFDMLTYCEHIWFYARLKG EKHVKAEMEGMALDVGLPSSKLKSKTSQL GGMQRKLSVALAFVGGSKVVILDEPTAGV YSRRGIWELLLKYPQGGTSKLVBLCTGYGYU GGRMANSHGKLCVGSSLFLKNOLGTGYYU FRRGIWELLLKYPQGGTRILSTTHMDEAD GDRAIISHGKLCCVGSSLFLKNOLGTGYYU FRRGIWELLLKYPQGTTILSTTHMDEAD							
976 2326 A 8257 298 7086 GNMACWPQLRLLLWKNLTFRRRQTCQLL VAWPLFIFLILISVRLSYPPYEQHECHPPNK PSAGTLPWVQGIICNANNPCRYPTPGEAP VGMFNKSIVARLFSDARRLLLYSQKDTSMI MRKVLRTLQQIKKSSSNLKLQDFLVDNETI GFLYHNLSLPKSTVDKMLRADVILHKVFLL YQUHLTSLCNSKSEEMIQLGDQEVSELC REKLAAAERVLRSNMDILKPILRTLNSTSPI KELAEATKTLLHSLGTLAQELFSMRSWSDI OEVMFLTNYSSSSTOTYQAVSRIVCGHP GGLKIKSLNWYEDNNYKALFGGNGTEEDA TIFYDNSTTPYCNDLMKNLESSPLSRIIWKA PLLLVGKILYTPDTPATRQVMAEVNKTFQEI VFHDLEGMWEELSPKIWTFMENSQEMDLA AKHPEDVQSSNGSVYTWREAFNETNQAIR RFMECVNLNKLEPIATEVWLNKSMELLDI KFWAGIVFTGITPGSIELPHHVKYKIRMGIL VERTNKKDGYWDDGPRADPFEDMRYVW FAYLQDVVEQAIRVLTGTEEKKTGYYMQQ YPCYVDDIFLRVMSRSMPLFMTLAWIYSV IIKGIVYEKEARLKETMRIMGLDNSILWFSV SSLIPLLVSAGLLVVILKLGNLLPYSDPSVV FLSVFAVVTILCCFLISTLFSRANLAAACGG YFTLYLPYVLCVAWQDVVGFTLKIFASLLS VAFGFGCEYFALFEEQGIGVQWDNLFESPY EDGFNLTTSVSMMLFDTFLYGVMTWYIEA PGQYGIPRPWYFPCTKSYWFGEESDEKSHP NQKRISEICMEEEPTHLKLGVSIQNLVKVYI GMKVAVDGLAINFYEGGITSFLGHNGAGK TTMSILTGLFPPTSGTAYILGKDIRSEMSTIR NLGVCPQHNVLFDMLTVEEHIWFYARLKG GKHVKAEMEQMALDVGIPSSKLKSKTSQL GGMQRKLSVALAFVGGGSKVVILDEPTAGV YSRGIWELLLKYRQGRTILLSTHHMDEAD GDRAIISHSHCLCCVGSSLFLKKNOLGTGYYL	1 1						PYWCHHCGKTFCSKSNLSKHORVHTGFGFA
976 2326 A 8257 298 7086 GNMACWPQLRLLLWKNLTFRRQTCQILL VAWPLFIFLLISVRLSYPPYEQHECHFPNK PSAGTLPWVQGIICNANNPCRYPTPGEAP VGNFNKSIVARLFSDARRLLLYSQKDTSM MRKVLRTLQQIKKSSSNLKLQDFLVNNETI GFLYHNLSLFSTYDKMLRADVILHKVFLL YQLHLTSLCNGSKSEEMIQLGDQEVSELCC REKLAAAERVLRSNMDILKPILRTLNTSTSPI KELAEATKILHSLGTLAQGLFSMRSWIVCGHP GGLKIKSLNWYEDNYKALFGGNGTEEDA TFYDNSTTPYCNDLMKNLESSPLSRIEWA PLLVGKILYTPDTPATRQVMAEVNKTFQEI VFHDLEGMWEELSPKIWTFMENSQEMDLV MLLDSRDNHPWEQQLISVLGUMTAQDIVA AKHPEDVQSSNGSVYTWREAFNETNQAIR RFMECVNLNKLEPIATEVWLINKSMELLDS KFWAGIVFTGFGSIELPHNYKYKRMGIE VERTNKIKDGYWDPGPRADPFEDMRYVW FAYLQDVVEQAIRVLTGTEKKTGYYMQQ YPCYVDDIFLRVMSRSMPLFMTLAWIYS SSLIPLLVSAGLLVVYLLKIGNLLPYSDPSVV IIKGIVYEKEARLKETMRIMGLDNSILWFSV SSLIPLLVSAGLLVVILKLGNLLPYSDPSVV FLSVFAVVTILQCFLISTLFSRANLAAACGG YFTLYLPYVLCVAWQDYVGFTLKIFASLLS VAFGFGCEYFALFEQGIGVQWDNLFESPV EDGFNLTTSVSMMLFDTFLYGVMTWYIEA PGQYGIPRPWYFPCTKSYWFGEESDEKSHP NQKRISEICMEEETHLKLGVSIGNLVKVYI GMKVAVDGLALNFYEGGITSFLGHNGAGK TTMSILTGLFPPTSGTAYILGKDRSEMSTIR NLGVCPQHNVLFIKLKGVILGKDRSEMSTIR NLGVCPQHNVLFMLTVEEHWFYARLKG EKHYKAEMBQMALDVGIPSSKLKSKTSGL GGMQRKLSVALAFVGGSKVVILDEPTAGV YSRGIWELLLKYRQGRTILLSTHHMDEAD GDRIAIISHGKLCCVGSSLFLKNOLGTGYYI	1 1						P
VAWPLFIFLILISVRLSYPPYEGHECHFPNK PSAGTLPWVQGIICNANNFCRYPTPGEAP VGNFNKSIVARLFSDARRLLLYSQKDTSMI MRKVLRTLQQIKKSSSNLKLQDFLVDNETI GFLYHNLSLPKSTVDKMLRADVILHKVFLI YQLHLTSLCNGSKSEEMIQLGDQEVSELCC REKLAAAERVLRSNMDILKPILRTLNSTSPI KELAEATKTLLHSLGTLAQELFSMRSWSDI QEVMFLTNVNSSSSTQIYQAVSRIVCGHPP GGLKIKSLNWYEDNNYKALFGGNGTEEDJ TFYDNSTTPYCNDLMKNLESSPLSRIIWKA PLLVGKILYTPDTARQVMAEVNKTFQEI VFHDLEGMWEELSPKIWTTMENSQEMDLV MLDSRDNDHFWEQQLOXLDWTAQDIVA AKHPEDVQSSNGSVYTWREAFNETNQAIR RFMECVNLNKLEPIATEVWLINKSMELLDI KFWAGIVFTGITPGSIELPHHVKYKIRMGIL VERTNKIKDGYWDPGPRADPFEDMRYVW FAYLQDVVEQAIRVLTGTEKKTGVYMQQ YPCYVDDIFLRVMSRSMPLFMTLAWIYSV IIKGIVYEKEARLKETTMRIMGLDNSILWFSS SILPLLVSAGLLVVILKLGNLLPYSDPSVV FLSVFAVVTILQCFLISTLFSRANLAAACGG YFTLYLPYVLCVAWQDYVGFTLKIFASLLS VAFGFGCEYFALFEQGIGVOWDNLFESPN EDGFNLTTSVSMMLFDTFLYGVMTWYIEA PGQYGIPRPWYFFCTKSYWFGEESDEKSHP NQKRISEICMEEPTHLKLGVSIQNLVKVYI GMKVAVDGLALNFYEGGITSFLGHNGAGK TTMSILTGLFPPTSGTAYILGKDIRSEMSTIR NLGVCPQHNVLFMLTVEEHIWFYARLKG EKHVKAEMEQMALDVGLPSSKLKSKTSQL GGMQRKLSVALAFVGGGKVVILDEPTAGV YSRRGIWELLLKYRQGSKVVIDEPTAGV YSRRGIWELLLKYRQGSKVVIDEPTAGV YSRRGIWELLLKYRQGRTIILSTHHMDEAD GDRIAIISHGKLCCVGSSLFLKNOLGTGYYI	976	2326	Α	8257	298	7086	
PSAGTLPWVQGIICNANNPCFRYPTPGEAP VGNFNKSIVARLFSDARRLLLYSQKDTSM MRKVLKTLQQIKKSSSNLKLQDFLVDNETI GFLYHNLSLPKSTVDKMLRADVILHKVFLI YQLHLTSLCNGSKSEEMIQLGDQEVSELCC REKLAAAERVLRSNMDJILKPILRTLNSTSPI KELAEATKTLLHSLGTLAQELFSMRSWSDI QEVMFLTNNNSSSSTQIYQAVSRIVCGHPP GGLKIKSLWYEDNNYKALFGGNGTEEDA TFYDNSTTPYCNDLMKNLESSPLSRIIWKA PLLVGKILYTPDTTATRQVMAEVNKTFQEI VFHDLEGMWEELSPKLWYFMENSQEMDLY MLLDSRDNDHFWEQQLDGLDWTAQDIVA AKHPEDVQSSNGSYYTWREAFNETNQAIR RFNECVOLNKLEPIATEVWLINKSMELLDI KFWAGIVFTGITPGSIELPHHVKYKIRMGIL VERTNKIKDGYWDPGPRADPFEDMRYVW FAYLQDVVEQAIRVLLTGTEKKTGVYMQQ YPCYVDDIFLRVMSRSMPLFMTLAWIYSV. IIKGIVYEKEARLKETMRIMGLDNSILWFSSY SSLIPLLVSAGLLVVILKLGNLLPYSDPSVV FLSVFAVVTILQCFLISTLFSRANLAAACGG YFTLYLPYVLCVAWQDYVGFTLKIFASLLS VAFGFGCEYFALFEEQGIGVQWDNLFESPV EDGFNLTTSVSMMLFDTFLYGVMTWYIEA PGQYGIPRPWYPPCTKSYWFGGESDEKSHP NQKRISEICMEEEPTHLKLGVSIQNLVKVYI GMKVAVDGLALNFYEQGITSFLGHNGAGK TTMSILTGLFPPTSGTAYILGKDRSEMSTIR NLGVCPQHNVLFDMLTVEEHIWFYARLKG EKHVKAEMEQMALDVGLPSSKLKSKTSQL GGMQRKLSVALAFVGGSKVVILDEPTAGV YSRRGIWELLLKYRQGRTILLSTHHMDEAD GDRAIISHGKLCCVGSSLFLKNOLGTGYYIL	1						
VGNFNKSIVARLFSDARRLLLYSQKDTSMI MRKVLRTLQQIKKSSSNIKLQDFLVDNETI GFLYHNJSLPKSTVDKMLRADVILHKVFL YQLHLTSLCNGSKSEEMIQLGDQEVSELCC REKLAAAERVILRSSMDILKPILRTLNSTSPJ KELAEATKTILHSLGTLAQELFSMRSWSD QEVMFLTNVNSSSSSTQIYQAVSRIVCGHP GGLKIKSLNWYEDNNYKALFGGNGTEEDA TFYDNSTTPYCNDLMKNLESSPLSRIIWKA PLLVGKILYTPDTPATRQVMAEVNKTFQEI VFHDLEGMWEELSPKIWTFMENSQEMDLA MLLDSRDNDHFWEQQLDGLDWTAQDIVA AKHPEDVQSSNGSVYTWREAFNETNQAIR RFMECVNLNKLEPIATEVWLINKSMELLDH KFWAGIVFTGITPGSIELPHHVKYKRIMGILD VERTNKIKDGYWDPGPRADPFEDMRYVW FAYLQDVVEQAIIRVLTGTEKKTGVYMQQ YPCYVDDIFLRVMSRSMPLFMTLAWIYSVA IIKGIVYEKEARLKETMRIMGLDNSILWFSV SSLIPLLVSAGLLVVILKLGNLLPYSDPSVV FLSVFAVVTILQCFLISTLFSRANLAAACGG YFTLYLPYVLCVAWQDYVGFTLKIFASLLS VARGFGCEYFALFEEQGIGVQWDNLFESSPV EDGFNLTTSVSMMLFDTFLYGVMTWYIEA PGQYGIPRPWYFPCTKSYWFGEESDEKSHP NOKRISEICMEEPFTHLKLGVSIQNLKVYY GMKVAVDGLALNFYEGQITSFLGHNGAGK TTMSILTGLFPPTSGTAYILGKDIRSEMSTIR NLGVCPQHNVLFDMLTVEEHIWFYARLKG EKHVKAEMEQMALDVGLPSSKLKSKTSQL GGMQRKLSVALAFYGGSKVVILDEPTAGV YSRRGIWELLLKYRQGRTIILSTHHMDEAD GORMISHGKLCCVGSSLFLKNOLGTGYYY	1 1						
MRKVLRTLQQIKKSSSNLKLQDFLVDNETI GFLYHNLSLPKSTVDKMLRADVILHKVFLI YQLHLTSLCNGSKSEEMIQLGDQEVSELCC REKLAAAERVLRSNMDILKPILRTLNSTSPI KELAEATKTLLHSLGTLAQELFSMRSWSDI QEVMFLTNVNSSSSSTQIYQAVSRIVCGHP GGLKIKSLNWYEDNNYKALFGGNGTEEDA TFYDNSTTPYCNDLMKNLESSPLSRIIWKA PLLVGKILYTPDTPATRQVMAEVNKTFQEI VFHDLEGMWEELSPKLWTFMENSQEMDLN MLLDSRDNDHFWEQQLDGLDWTAQDIVA AKHPEDVGSNGSVYTWREAFNETNQAIR RFMECVNLNKLEPIATEVWLINKSMELLDF KFWAGIVFTGITTGSIELPHHVKYKIRMGIL VERTNKIKDGYWDPGPRADPFEDMRYVW FAYLQDVVEQAIIRVLTGTEKKTGVYMQQ YPCYVDDIFLRVMSRSMPLFMTLAWIYSV, IIKGIVYEKEARLKETMRIMGLDNSILWFSV SSLIPLLVSAGLLVVILKLGNLLPYSDPSVV FLSVFAVVTILQCFLISTLFSRANLAAACGG YFFLYLPVVLCVAWQDVYGFTLKIFASLLS VAGGFGCEYFALFEEQGIGVQWDNLFESFY EDGFNLTTSVSMMLFDTFLYGVMTWYIEA PGQYGIPRPWYFPCTKSYWFGEESDEKSHP NQKRISCIMEEDFTHLKLGVSIQOLVKVYI GMKVAVDGLALNFYEGQITSFLGHNGAGK TTMSILTGLFPPTSGTAYILGKDIRSEMSTIR NLGVCPQHNVLFDMLTVEEHIWFYARLKG GMKVAVDGLALNFYEGQITSFLGHNGAGK TTMSILTGLFPPTSGTAYILGKDIRSEMSTIR NLGVCPQHNVLFDMLTVEEHIWFYARLKG GMKVAVDGLALNFYEGQITSFLGHNGAGK TTMSILTGLFPPTSGTAYILGKDIRSEMSTIR NLGVCPQHNVLFDMLTVEEHIWFYARLKG GKVAVDGLALNFYGGGSPEKVILDEPTAGV YSRRGIWELLLKYRQGRTIILSTHHMDEAD GGRAILSHGKLCCVGSSLFLKNOLGTGFYYL	1						
GFLYHNLSLPKSTVDKMLRADVILHKVFL YQLHLTSLCNGSKSEEMIQLGDQEVSELCC REKLAAAERVLRSNMDILKPILRSTLNSTSPJ KELAEATKTLLHSLGTLAQELFSMRSWSD QEVMFLTNVNSSSSSTQIYQAVSRIVCGHP GGLKIKSLNWYEDNNYKALFGGNGTEEDA TFYDNSTTPYCNDLMKNLESSPLSRIIWKA PLLVGKILYTPDTPATRQVMAEVNKTFQEI VFHDLEGMWEELSPKLWTFMENSQEMDLN MLDSBNDHFWEQQLDGLDWTAQDIVA AKHPEDVQSSNGSVYTWREAFNETNQAIR RFMECVNLNKLEPIATEVWLINKSMELLDI KFYMGIVFTGITPGSIELPHHVKYKIRMGID VERTNKIKDGYWDPGPRADPFEDMRYVW FAYLQDVVEQAIRVLTGTEKKTGVYMQQ YPCYVDDIFLRVMSRSMPLFMTLAWIYSV. IIKGIVYEKEARLKETMRIMGLDNSILWFSSY SSLIPLLVSAGLLVVILKLGNLLPYSDPSVV FLSVFAVVTILQCFLISTLFSRANLAAACGG YFTLYPYLCVAWQDYVGFTLKIFASLLS VAFGFGCEYFALFEEQGIGVQWDNLFESPV EDGFNLTTSVSMMLFDTFLYGVMTWYIEA PGQYGIPPPWYFPCTKSYWFGEESDEKSHP NQKRISEICMEEPFTHLKLGVSIQNLVKVYI GMKVAVDGLALNFYEGQITSFLGHNGAGK TTMSILTGLFPPTSGTAYILGKDIRSEMSTIR NLGVCPQHNVLFDMLTVEEHHWYARLKG EKHVKAEMEQMALDVGLPSSKLKSKTSQL GGMQRKLSVALAFVGGSKVVILDEPTAGV YSRRGIWELLLKYRQGRTIILSTHHMDEAD GGRAILSHGKLCCVGSSLFLKNOLGTGYYL	1 1						
YQLHLTSLCNGSKSEEMIQLGDQEVSELCC REKLAAAERVLRSNMDILKPILRTILNTSTSPI KELAEATKTLLHSLGTLAQELFSMRSWSDI QEVMFLTNVNSSSSTQIYQAVSRIVCGHP GGLKIKSLNWYEDNNYKALFGGNGTEEDA TFYDNSTTPYCNDLMKNLESSPLSRIIWKA PLLVGKILYTPDTPATRQVMAEVNKTFQEI VFHDLEGMWEELSPKIWTFMENSQEMDLN MLLDSRDNDHFWEQQLDGLDWTAQDIVA AKHPEDVQSSNGSVYTWREAFNETNQAIR RFMECVNLNKLEPIATEVWLINKSMELLDI KFWAGIVFTGITPGSIELPHHVKYKIRMGID VERTNKIKDGYWDPGPRADPFEDMRYVW FAYLQDVVEQAIIRVLTGTEKKTGVYMQQ YPCYVDDIFLRVMSRSMPLFMTLAWIYSV. IIKGIVYEKEARLKETMRIMGLDNSILWFSV SSLIPLLVSAGLLVVILKLGNLLPYSDPSVV FLSVFAVVTILQCFLISTLFSRANLAAACGG YFTLYLPYVLCVAWQDYVGFTLKIFASLLS VAFGFGCEYFALFEQGIGVQWDNLFESPV EDGFNLTTSVSMMLFDTFLYGVMTWYIEA PGQGIPRPWYPPCTKSYWFGEESDEKSHP NQKRISEICMEEEPTHLKLGVSIQNLVKVYI GMKVAVDGLALNFYEGQITSFLGHNGAGK TTMSILTGLFPPTSGTAYILGKDIRSEMSTIR NLGVCPQHNVLFDMLTVEEHIWFYARLKG EKHVKAEMEQMALDVGLPSSKLKSKTSQL GGMQRKLSVALAFVGGSKVVILDEPTAGV YSRRGIWELLLKYRQGRTIILSTHHMDEAD GORNAINISHGKLCCVGSSLFLKNOLGTGYYI	1 1			1	i		
REKLAAAERVLRSNMDILKPILRTLNSTSPI KELAEATKTLHSLGTLAQELFSMRSWSD) QEVMFLTNVNSSSSTQIYQAVSRIVCGHP GGLKIKSLNWYEDNNYKALFGGNGTEEDA TFYDNSTTPYCNDLMKNLESSPLSRIIWKA PLLVGKILYTPDTPATRQVMAEVNKTFQEI VFHDLEGMWEELSPKIWTFMENSQEMDLY MLDSRDNDHFWEQQLDGLDWTAQDIVA AKHPEDVQSSNGSVYTWREAFNETNQAIR RFMECVNLNKLEPIATEVWLINKSMELLDI KFWAGIVFTGITPGSIELPHHVKYKIRMGIL VERTNKIKDGYWDPGPRADPFEDMRYVW FAYLQDVVEQAIIRVLTGTEKKTGYYMQQ YPCYVDDIFLRVMSRSMPLFMTLAWIYSV. IIKGIVYEKEARLKETMRIMGLDNSILWFSV SSLIPLLVSAGLLVVILKLGNLLPYSDPSVV FLSVFAVVTILQCFLISTLFSRANLAAACGG YFITLYLPYVLCVAWQDYVGFTLKIFASILS VAFGFGCEYFALFEQGIGVQWDNLFESPV EDGFNLTTSVSMMLFDTFLYGVMTWYIEA PGQYGPPWYPPCTKSYWFGEESDEKSHP NQKRISEICMEEEPTHLKLGVSIQNLVKVYI GMKVAVDGLALNFYEGQITSFLGHNGAGK TTMSILTGLFPTSGTAYILGKDIRSEMSTIR NLGVCPQHNVLFDMLTVEEHIWFYARLKG EKHVKAEMEQMALDVGLPSSKLKSKTSQL GGMQRKLSVALAFVGGSKVVILDEPTAGV YSRRGIWELLLKYRQGRTIILSTHHMDEAD GORIAIISHGKLCCVGSSLFLKNOLGTGYYI					ļ		YOLHLTSLCNGSKSEEMIOLGDOEVSELCGLP
KELAEATKTLLHSLGTLAQELFSMRSWSDI QEVMFLTNVNSSSSTQIYQAVSRIVCGHPP GGLKIKSLNWYEDNNYKALFGGNGTEEDA TFYDNSTTPYCNDLMKNLESSPLSRIIWKA PLLVGKILYTPDTPATRQVMAEVNKTFQEI VFHDLEGMWEELSPKIWTFMENSQEMDLV MLDSSRDNDHFWEQQLDGLDWTAQDIVA AKHPEDVQSSNGSVYTWREAFNETNQAIR RFMECVNLNKLEPIATEVWLINKSMELLDI KFWAGIVFTGITTGSIELPHHVKYKIRMGIL VFRTNKIKDGYWDPGPRADPFEDMRYVW FAYLQDVVEQAIIRVLTGTEKKTGVYMQQ YPCYVDDIFLRVMSRSMPLFMTLAWIYSV IIKGIVYEKEARLKETMRIMGLDNSILWFSV SSLIPLLVSAGLLVVILKLGNLLPYSDPSVV FLSVFAVVTILQCFLISTLFSRANLAAACGG YFTLYLPYVLCVAWQDYVGFTLKIFASLLS VAFGFGCEVFALFEEQGIGVQWDNLFESPV EDGFNLTTSVSMMLFDTFLYGVMTWYIEA PGQYGIPRPWYFPCTKSYWFGEESDEKSHP NQKRISEICMEEEPTHLKLGVSIQNLVKVYI GMKVAVDGLALNFYEGQITSFLGHNGAGK TTMSILTGLFPPTSGTAYILGKDIRSEMSTIR NLGVCPQHNVLFDMLTVEEHIWFYARLKG EKHVKAEMEQMALDVGLPSSKLKSKTSQL GGMQRKLSVALAFVGGSKVVLDEPTAGV YSRRGIWELLLKYRQGRTIILSTHHMDEAD GDRIAIISHGKLCCVGSSLFLKNOLGTGYYI							REKLAAAERVLRSNMDILKPILRTLNSTSPEPS
QEVMFLTNVNSSSSSTQIYQAVSRIVCGHP GGLKKSLNW YEDNNYKALFGGNGTEEDA TFYDNSTTPYCNDLMKNLESSPLSRIIWKA PLLVGKILYTPDTPATRQVMAEVNKTFQEI VFHDLEGMWEELSPKIWTFMENSQEMDLY MLLDSRDNDHFWEQQLDGLDWTAQDIVA AKHPEDVQSSNGSVYTWREAFNETNQAIR RFFMECVNLNKLEPIATEVWLINKSMELLDI KFWAGIVFTGITPGSIELPHHVKYKIRMGID VERTNKIKDGYWDPGPRADPFEDMRYVW FAYLQDVVEQAIIRVLTGTEKKTGVYMQQ YPCYVDDIFLRVMSRSMPLFMTLAWIYSV IIKGIVYEKEARLKETMRIMGLDNSILWFSY SSLIPLLVSAGLLVVILKLGNLLPYSDPSVV FLSVFAVVTILQCFLISTLFSRANLAAACGG YFTLYLPYVLCVAWQDYVGFTLKIFASLLS VAFGFGCEYFALFEEQGIGVQWDNLFESPV EDGFNLTTSVSMMLFDTFLYGVMTWYIEA PGQYGIPRPWYFPCTKSYWFGEESDEKSHP NQKRISEICMEEPTHLKLGVSIQNLVKVYJ GMKVAVDGLALNFYEGQITSFLGHNGAGK TTTMSILTGLFPPTSGTAYILGKDRSEMSTIR NLGVCPQHNVLFDMLTVEEHIWFYARLKG EKHVKAEMEQMALDVGLPSSKLKSKTSQL GGMQRKLSVALAFVGGSKVVILDEPTAGV YSRRGIWELLLKYRQGRTIILSTHHMDEAD GDRIAIISHGKLCCVGSSLFLKNOLGTGYYI]						KELAEATKTLLHSLGTLAOELFSMRSWSDMR
GGLKIKSLNWYEDNNYKALFGGNGTEEDA TFYDNSTTPYCNDLMKNIESSPLSRIIWKA PLLVGKILYTPDTPATRQVMAEVNKTFQEI VFHDLEGMWEELSPKIWTFMENSQEMDLY MLLDSRDNDHFWEQQLDGLDWTAQDIVA AKHPEDVQSSNGSVYTWREAFNETNQAIR RFMECVNLNKLEPIATEVWLINKSMELDI KFFWAGIVFTGITPGSIELPHHVKYKIRMGIL VERTNKIKDGYWDPGPRADPFEDMRYVW FAYLQDVVEQAIIRVLTGTEKKTGVYMQQ YPCYVDDIFLRVMSRSMPLFMTLAWIYSV. IIKGIVYEKEARLKETMRIMGLDNSILWFSV SSLIPLLVSAGLLVVILKLGNLLPYSDPSVV FLSVFAVVTILQCFLISTLFSRANLAAACGG YFTLYLPYVLCVAWQDYVGFTLKIFASLLS VAFGFGCEYFALFEEQGIGVQWDNLFESPV EDGFNLTTSVSMMLFDTFLYGVMTWYIEA PGQYGIPRPWYFPCTKSYWFGEESDEKSHP NQKRISEICMEEPTHLKLGVSIQNLVKVYI GMKVAVDGLALNFYEGQITSFLGHNGAGK TTTMSILTGLFPPTSGTAYILGKDRSEMSTIR NLGVCPQHNVLFDMLTVEEHIWFYARLKG EKHVKAEMEQMALDVGLPSSKLKSKTSQL GGMQRKLSVALAFVGGSKVVILDEPTAGV YSRRGIWELLIKYRQGRTIILSTHHMDEAD GDRIAIISHGKLCCVGSSLFLKNOLGTGYYI	1 1	ĺ			ĺ	·	
TFYDNSTTPYCNDLMKNLESSPLSRIIWKA PLLVGKILYTPDTPATRQVMAEVNKTFQEI VFHDLEGMWEELSPKIWTFMENSQEMDLV MLLDSRDNDHFWEQQLDGLDWTAQDIVA AKHPEDVQSSNGSVYTWREAFNETNQAIR RFMECVNLNKLEPIATEVWLINKSMELLDI KFWAGIVFTGITPGSIELPHHVKYKIRMGID VERTNKIKDGYWDPGPRADPFEDMRYVW FAYLQDVVEQAIIRVLTGTEKKTGVYMQQ YPCYVDDIFLRVMSRSMPLFMTLAWIYSV. IIKGIVYEKEARLKETMRIMGLDNSILWFSV SSLIPLLVSAGLLVVILKLGNLLPYSDPSVV FLSVFAVVTILQCFLISTLFSRANLAAACGG YFTLYLPYVLCVAWQDYVGFTLKIFASLLS VAFGFGCEYFALFEEQGIGVQWDNLFESPV EDGFNLTTSVSMMLFDTFLYGVMTWYIEA PGQYGIPRPWYFPCTKSYWFGEESDEKSHP NQKRISEICMEEEPTHLKLGVSIQNLVKVYI GMKVAVDGLALNFYEGQITSFLGHNGAGK TTMSILTGLFPPTSGTAYILGKDIRSEMSTIR NLGVCPQHNVLFDMLTVEEHIWFYARLKG EKHVKAEMEQMALDVGLPSSKLVSKTSQL GGMQKLSVALAFVGGSKVVILDEPTAGV YSRRGIWELLLKYRQGRTIILSTHHMDEAD GDRIAIISHGKLCCVGSSLFLKNOLGTGYYL	1 1				1		
PLLVGKILYTPDTPATRQVMAEVNKTFQEI VFHDLEGMWEELSPKLWTFMENSQEMDLV MLLDSRDNDHFWEQQLDGLDWTAQDIVA AKHPEDVQSSNGSVYTWREAFNETNQAIR RFMECVNLNKLEPIATEVWLINKSMELLDE KFWAGIVFTGITPGSIELPHHVKYKIRMGID VERTNKIKDGYWDPGPRADPFEDMRYVW FAYLQDVVEQAIIRVLTGTEKKTGVYMQQ YPCYVDDIFLRVMSRSMPLFMTLAWIYSVV IIKGIVYEKEARLKETMRIMGLDNSIL WFSV SSLIPLLVSAGLLVVILKLGNLLPYSDPSVV FLSVFAVVTILQCFLISTLFSRANLAAACGG YFTLYLPYVLCVAWQDYVGFTLKIFASLLS VAFGFGCEYFALFEEQGIGVQWDNLFESPV EDGFNLTTSVSMMLFDTFLYGVMTWYIEA PGQYGIPRPWYFPCTKSYWFGEESDEKSHP NQKRISEICMEEEPTHLKLGVSIQNLVKVYI GMKVAVDGLALNFYEGQITSFLGHNGAGK TTMSILTGLFPPTSGTAYILGKDIRSEMSTIR NLGVCPQHNVLFDMLTVEEHIWFYARLKG EKHVKAEMEQMALDVGLPSSKLKSKTSQL GGMQRKLSVALAFVGGSKVVILDEPTAGV YSRRGIWELLLKYRQGRTIILSTHHMDEAD GDRIAIISHGKLCCVGSSLFLKNOLGTGYYI	1 1						
VFHDLEGMWEELSPKIWTFMENSQEMDLY MLLDSRDNDHFWEQQLDGLDWTAQDIVA AKHPEDVQSSNGSVYTWREAFNETNQAIR RFMECVNLNKLEPIATEVWLINKSMELLDH KFWAGIVFTGITPGSIELPHHVKYKIRMGIU VERTNKIKDGYWDPGPRADPFEDMRYVW FAYLQDVVEQAIIRVLTGTEKKTGVYMQQ YPCYVDDIFLRVMSRSMPLFMTLAWIYSV. IIKGIVYEKEARLKETMRIMGLDNSIL WFSV SSLIPLLVSAGLLVVILKLGNLLPYSDPSVV FLSVFAVVTILQCFLISTLFSRANLAAACGG YFTLYLPYVLCVAWQDYVGFTLKIFASLLS VAFGFGCEYFALFEEQGIGVQWDNLFESPV EDGFNLTTSVSMMLFDTFLYGVMTWYIEA PGQYGIPRPWYFPCTKSYWFGEESDEKSHP NQKRISEICMEEEPTHLKLGVSIQNLVKVYI GMKVAVDGLALNFYEGQITSFLGHNGAGK TTMSILTGLFPPTSGTAYILGKDIRSEMSTIR NLGVCPQHNVLFDMLTVEEHIWFYARLKG EKHVKAEMEQMALDVGLPSSKLKSKTSQL GGMRKLSVALAFVGGSKVVILDEPTAGV YSRRGIWELLLKYRQGRTIILSTHHMDEAD GDRIAIISHGKLCCVGSSLFLKNOLGTGYYI	}						PLLVGKILYTPDTPATRQVMAEVNKTFOELA
AKHPEDVQSSNGSVYTWREAFNETNQAIR RFMECVNLNKLEPIATEVWLINKSMELLDE KFWAGIVFTGITPGSIELPHHVKYKIRMGIL VERTNKIKDGYWDPGPRADPFEDMRYVW FAYLQDVVEQAIIRVLTGTEKKTGVYMQQ YPCYVDDIFLRVMSRSMPLFMTLAWIYSV. IIKGIVYEKEARLKETMRIMGLDNSILWFSV SSLIPLLVSAGLLVVILKLGNLLPYSDPSVV FLSVFAVVTILQCFLISTLFSRANLAAACGG YFTLYLPYVLCVAWQDYVGFTLKIFASLLS VAFGFGCEYFALFEEQGIQVQWDNLFESPV EDGFNLTTSVSMMLFDTFLYGVMTWYIEA PGQYGIPRPWYFPCTKSYWFGEESDEKSHP NQKRISEICMEEEPTHLKLGVSIQNLVKVYI GMKVAVDGLALNFYEGQITSFLGHNGAGK TTMSILTGLFPPTSGTAYILGKDIRSEMSTIR NLGVCPQHNVLFDMLTVEEHIWFYARLKG EKHVKAEMEQMALDVGLPSSKLKSKTSQL GGMQRKLSVALAFVGGSKVVILDEPTAGV YSRRGIWELLLKYRQGRTIILSTHHMDEAD GDRIAIISHGKLCCVGSSLFLKNOLGTGYYL	<u> </u>						VFHDLEGMWEELSPKIWTFMENSQEMDLVR
AKHPEDVQSSNGSVYTWREAFNETNQAIR RFMECVNLNKLEPIATEVWLNKSMELLDE KFWAGIVFTGITPGSIELPHHVKYKIRMGIL VERTNKIKDGYWDPGPRADPFEDMRYVW FAYLQDVVEQAIIRVLTGTEKKTGVYMQQ YPCYVDDIFLRVMSRSMPLFMTLAWIYSV. IIKGIVYEKEARLKETMRIMGLDNSILWFSV SSLIPLLVSAGLLVVILKLGNLLPYSDPSVV FLSVFAVVTILQCFLISTLFSRANLAAACGG YFTLYLPYVLCVAWQDYVGFTLKIFASLLS VAFGFGCEYFALFEEQGIQVQWDNLFESPV EDGFNLTTSVSMMLFDTFLYGVMTWYIEA PGQYGIPRPWYFPCTKSYWFGEESDEKSHP NQKRISEICMEEEPTHLKLGVSIQNLVKVYI GMKVAVDGLALNFYEGQITSFLGHNGAGK TTMSILTGLFPPTSGTAYILGKDIRSEMSTIR NLGVCPQHNVLFDMLTVEEHIWFYARLKG EKHVKAEMEQMALDVGLPSKLKSKTSQL GGMQRKLSVALAFVGGSKVVILDEPTAGV YSRRGIWELLLKYRQGRTIILSTHHMDEAD GDRIAIISHGKLCCVGSSLFLKNOLGTGYYL							MLLDSRDNDHFWEQQLDGLDWTAQDIVAFL
RFMECVNLNKLEPIATEVWLINKSMELLDI KFWAGIVFTGITPGSIELPHHVKYKIRMGID VERTNKIKDGYWDPGPRADPFEDMRYVW FAYLQDVVEQAIIRVLTGTEKKTGVYMQQ YPCYVDDIFLRVMSRSMPLFMTLAWIYSV IIKGIVYEKEARLKETMRIMGLDNSILWFSV SSLIPLLVSAGLLVVILKLGNLLPYSDPSVV FLSVFAVVTILQCFLISTLFSRANLAAACGG YFTLYLPYVLCVAWQDYVGFTLKIFASLLS VAFGFGCEYFALFEEQGIGVQWDNLFESPV EDGFNLTTSVSMMLFDTFLYGVMTWYIEA PGQYGIPRPWYFPCTKSYWFGEESDEKSHP NQKRISEICMEEEPTHLKLGVSIQNLVKVYI GMKVAVDGLALNFYEGQITSFLGHNGAGK TTMSILTGLFPPTSGTAYILGKDRSEMSTIR NLGVCPQHNVLFDMLTVEEHIWFYARLKG EKHVKAEMEQMALDVGLPSSKLKSKTSQL GGMQRKLSVALAFVGGSKVVILDEPTAGV YSRRGIWELLLKYRQGRTIILSTHHMDEAD GDRIAIISHGKLCCVGSSLFLKNOLGTGYYL	1 1						AKHPEDVQSSNGSVYTWREAFNETNOAIRTIS
KFWAGIVFTGITPGSIELPHHVKYKIRMGID VERTNKIKDGYWDPGPRADPFEDMRYVW FAYLQDVVEQAIIRVLTGTEKKTGVYMQQ YPCYVDDIFLRVMSRSMPLFMTLAWIYSVY IIKGIVYEKEARLKETMRIMGLDNSILWFSV SSLIPLLVSAGLLVVILKLGNLLPYSDPSVV FLSVFAVVTILQCFLISTLFSRANLAAACGG YFTLYLPYVLCVAWQDYVGFTLKIFASLLS VAFGFGCEYFALFEEQGIGVQWDNLFESPV EDGFNLTTSVSMMLFDTFLYGVMTWYIEA PGQYGIPRPWYFPCTKSYWFGEESDEKSHP NQKRISEICMEEEPTHLKLGVSIQNLVKVYYI GMKVAVDGLALNFYEGQITSFLGHNGAGK TTMSILTGLFPPTSGTAYILGKDRSEMSTIR NLGVCPQHNVLFDMLTVEEHIWFYARLKG EKHVKAEMEQMALDVGLPSSKLKSKTSQL GGMQRKLSVALAFVGGSKVVILDEPTAGV YSRRGIWELLLKYRQGRTIILSTHHMDEAD GDRIAIISHGKLCCVGSSLFLKNOLGTGYYL	1 1	ľ	ł			i	RFMECVNLNKLEPIATEVWLINKSMELLDER
FAYLQDVVEQAIRVLTGTEKKTGVYMQQ YPCYVDDIFLRVMSRSMPLFMTLAWIYSV IIKGIVYEKEARLKETMRIMGLDNSILWFSV SSLIPLLVSAGLLVVILKLGNLLPYSDPSVV FLSVFAVVTILQCFLISTLFSRANLAAACGG YFTLYLPYVLCVAWQDYVGFTLKIFASLLS VAFGFGCEYFALFEEQGIGVQWDNLFESPV EDGFNLTTSVSMMLFDTFLYGVMTWYIEA PGQYGIPRPWYFPCTKSYWFGEESDEKSHP NQKRISEICMEEEPTHLKLGVSIQNLVKVYY GMKVAVDGLALNFYEGQITSFLGHNGAGK TTMSILTGLFPPTSGTAYILGKDIRSEMSTIR NLGVCPQHNVLFDMLTVEEHIWFYARLKG EKHVKAEMEQMALDVGLPSSKLKSKTSQL GGMQRKLSVALAFVGGSKVVILDEPTAGV YSRRGIWELLLKYRQGRTIILSTHHMDEAD GDRIAIISHGKLCCVGSSLFLKNOLGTGYYL			1			i	KFWAGIVFTGITPGSIELPHHVKYKIRMGIDN
YPCYVDDIFLRVMSRSMPLFMTLAWIYSVA IIKGIVYEKEARLKETMRIMGLDNSILWFSV SSLIPLLVSAGLLVVILKLGNLLPYSDPSVV FLSVFAVVTILQCFLISTLFSRANLAAACGG YFTLYLPYVLCVAWQDYVGFTLKIFASLLS VAFGFGCEYFALFEQGIGVQWDNLFESPV EDGFNLTTSVSMMLFDTFLYGVMTWYIEA PGQYGIPRPWYFPCTKSYWFGEESDEKSHP NQKRISEICMEEEPTHLKLGVSIQNLVKVYI GMKVAVDGLALNFYEGQITSFLGHNGAGK TTMSILTGLFPPTSGTAYILGKDIRSEMSTIR NLGVCPQHNVLFDMLTVEEHIWFYARLKG EKHVKAEMEQMALDVGLPSSKLKSKTSQL GGMQRKLSVALAFVGGSKVVILDEPTAGV YSRRGIWELLLKYRQGRTIILSTHHMDEAD GDRIAIISHGKLCCVGSSLFLKNOLGTGYYL	ł		- 1		1		VERTNKIKDGYWDPGPRADPFEDMRYVWGG
IIKGIVYEKEARLKETMRIMGLDNSILWFSV SSLIPLLVSAGLLVVILKLGNLLPYSDPSVV FLSVFAVVTILQCFLISTLFSRANLAAACGG YFTLYLPYVLCVAWQDYVGFTLKIFASLLS VAFGFGCEYFALFEEQGIQVQWDNLFESPV EDGFNLTTSVSMMLFDTFLYGVMTWYIEA PGQYGIPRPWYFPCTKSYWFGEESDEKSHP NQKRISEICMEEEPTHLKLGVSIQNLVKVYI GMKVAVDGLALNFYEGQITSFLGHNGAGK TTMSILTGLFPPTSGTAYILGKDIRSEMSTIR NLGVCPQHNVLFDMLTVEEHIWFYARLKG EKHVKAEMEQMALDVGLPSKLKSKTSQL GGMQRKLSVALAFVGGSKVVILDEPTAGV YSRRGIWELLLKYRQGRTIILSTHHMDEAD GDRIAIISHGKLCCVGSSLFLKNOLGTGYYL							FAYLQDVVEQAIIRVLTGTEKKTGVYMQQMP
SSLIPLLVSAGLLVVILKLGNLLPYSDPSVV FLSVFAVVTILQCFLISTLFSRANLAAACGG YFTLYLPYVLCVAWQDYVGFTLKIFASLLS VAFGFGCEYFALFEEQGIQVQWDNFESPV EDGFNLTTSVSMMLFDTFLYGVMTWYIEA PGQYGIPRPWYFFCTKSYWFGEESDEKSHP NQKRISEICMEEEPTHLKLGVSIQNLVKVYI GMKVAVDGLALNFYEGQITSFLGHNGAGK TTMSILTGLFPPTSGTAYILGKDIRSEMSTIR NLGVCPQHNVLFDMLTVEEHIWFYARLKG EKHVKAEMEQMALDVGLPSKLKSKTSQL GGMQRKLSVALAFVGGSKVVILDEPTAGV YSRRGIWELLLKYRQGRTIILSTHHMDEAD GDRIAIISHGKLCCVGSSLFLKNOLGTGYYL	1 . 1	1		ľ		ì	YPCYVDDIFLRVMSRSMPLFMTLAWIYSVAV
FLSVFAVVTILQCFLISTLFSRANLAAACGG YFTLYLPYVLCVAWQDYVGFTLKIFASLLS VAFGFGCEYFALFEEQGIGVQWDNLFESPV EDGFNLTTSVSMMLFDTFLYGVMTWYIEA PGQYGIPRPWYFFCTKSYWFGEESDEKSHP NQKRISEICMEEEPTHLKLGVSIQNLVKVYI GMKVAVDGLALNFYEGQITSFLGHNGAGK TTMSILTGLFPPTSGTAYILGKDIRSEMSTIR NLGVCPQHNVLFDMLTVEEHIWFYARLKG EKHVKAEMEQMALDVGLPSSKLKSKTSQL GGMQRKLSVALAFVGGSKVVILDEPTAGV YSRRGIWELLLKYRQGRTIILSTHHMDEAD GDRIAIISHGKLCCVGSSLFLKNOLGTGYYL	1		1	-	1		IIKGIVYEKEARLKETMRIMGLDNSILWFSWFI
YFTLYLPYVLCVAWQDYVGFTLKIFASLLS VAFGFGCEYFALFEEQGIGVQWDNLFESPV EDGFNLTTSVSMMLFDTFLYGVMTWYIEA PGQYGIPRPWYFFCTKSYWFGEESDEKSHP NQKRISEICMEEEPTHLKLGVSIQNLVKVYI GMKVAVDGLALNFYEGQITSFLGHNGAGK TTMSILTGLFPPTSGTAYILGKDIRSEMSTIR NLGVCPQHNVLFDMLTVEEHIWFYARLKG EKHVKAEMEQMALDVGLPSSKLKSKTSQL GGMQRKLSVALAFVGGSKVVILDEPTAGV YSRRGIWELLLKYRQGRTIILSTHHMDEAD GDRIAIISHGKLCCVGSSLFLKNOLGTGYYL		1		ł			SSLIPLLVSAGLLVVILKLGNLLPYSDPSVVFV
VAFGFGCEYFALFEEQGIGVQWDNLFESPV EDGFNLTTSVSMMLFDTFLYGVMTWYIEA PGQYGIPRPWYFPCTKSYWFGEESDEKSHP NQKRISEICMEEEPTHLKLGVSIQNLVKVYI GMKVAVDGLALNFYEGQITSFLGHNGAGK TTMSILTGLFPPTSGTAYILGKDIRSEMSTIR NLGVCPQHNVLFDMLTVEEHIWFYARLKG EKHVKAEMEQMALDVGLPSKLKSKTSQL GGMQRKLSVALAFVGGSKVVILDEPTAGV YSRRGIWELLLKYRQGRTIILSTHHMDEAD GDRIAIISHGKLCCVGSSLFLKNOLGTGYYL	1	j		j		ļ	FLSVFAVVTILQCFLISTLFSRANLAAACGGII
EDGFNLTTSVSMMLFDTFLYĞVMTWYIEA PGQYGIPRPWYFPCTKSYWFGEESDEKSHP NQKRISEICMEEEPTHLKLGVSIQNLVKVYY GMKVAVDGLALNFYEGQITSFLGHNGAGK TTMSILTGLFPPTSGTAYILGKDRSEMSTIR NLGVCPQHNVLFDMLTVEEHIWFYARLKG EKHVKAEMEQMALDVGLPSKLKSKTSQL GGMQRKLSVALAFVGGSKVVILDEPTAGV YSRRGIWELLLKYRQGRTIILSTHHMDEAD GDRIAIISHGKLCCVGSSLFLKNOLGTGYYL				-			YFTLYLPYVLCVAWQDYVGFTLKIFASLLSP
PGQYGIPRPWYFPCTKSYWFGEESDEKSHP NQKRISEICMEEEPTHLKLGVSIQNLVKVYI GMKVAVDGLALNFYEGQITSFLGHNGAGK TTMSILTGLFPPTSGTAYILGKDIRSEMSTIR NLGVCPQHNVLFDMLTVEEHIWFYARLKG EKHVKAEMEQMALDVGLPSKLKSKTSQL GGMQRKLSVALAFVGGSKVVLIDEPTAGV YSRRGIWELLLKYRQGRTIILSTHHMDEAD GDRIAIISHGKLCCVGSSLFLKNOLGTGYYL	1	l		1	İ		VAFGFGCEYFALFEEQGIGVQWDNLFESPVE
NQKRISEICMEEEPTHLKLGVSIQNLVKVYI GMKVAVDGLALNFYEGQITSFLGHNGAGK TTMSILTGLFPPTSGTAYILGKDIRSEMSTIR NLGVCPQHNVLFDMLTVEEHIWFYARLKG EKHVKAEMEQMALDVGLPSKLKSKTSQL GGMQRKLSVALAFVGGSKVVILDEPTAGV YSRRGIWELLLKYRQGRTIILSTHHMDEAD GDRIAIISHGKLCCVGSSLFLKNOLGTGYYL	1	l	1		1		EDGFNLTTSVSMMLFDTFLYGVMTWYIEAVF
GMKVAVDGLALNFYEGQITSFLGHNGAGK TTMSILTGLFPPTSGTAYILGKDIRSEMSTIR NLGVCPQHNVLFDMLTVEEHIWFYARLKG EKHVKAEMEQMALDVGLPSSKLKSKTSQL GGMQRKLSVALAFVGGSKVVILDEPTAGV YSRRGIWELLLKYRQGRTIILSTHHMDEAD GDRIAIISHGKLCCVGSSLFLKNOLGTGYYL	1	ł	ł		. 1		PGQYGIPRPWYFPCTKSYWFGEESDEKSHPGS
TTMSILTGLFPPTSGTAYILGKDIRSEMSTIR NLGVCPQHNVLFDMLTVEEHIWFYARLKG EKHVKAEMEQMALDVGLPSSKLKSKTSQL GGMQRKLSVALAFVGGSKVVILDEPTAGV YSRRGIWELLLKYRQGRTIILSTHHMDEAD GDRIAIISHGKLCCVGSSLFLKNOLGTGYYL			Ì		l		NQKRISEICMEEEPTHLKLGVSIQNLVKVYRD
NLGVCPQHNVLFDMLTVEEHIWFYARLKG EKHVKAEMEQMALDVGLPSSKLKSKTSQL GGMQRKLSVALAFVGGSKVVILDEPTAGV. YSRRGIWELLLKYRQGRTIILSTHHMDEAD GDRIAIISHGKLCCVGSSLFLKNOLGTGYYL		Ì	l		1		GMKVAVDGLALNFYEGQITSFLGHNGAGKT
EKHVKAEMEQMALDVGLPSSKLKSKTSQL GGMQRKLSVALAFVGGSKVVILDEPTAGV. YSRRGIWELLLKYRQGRTIILSTHHMDEAD GDRIAIISHGKLCCVGSSLFLKNOLGTGYYL	1		!		1		TTMSILTGLFPPTSGTAYILGKDIRSEMSTIRQ
GGMQRKLSVALAFVGGSKVVILDEPTAGV YSRRGIWELLLKYRQGRTIILSTHHMDEAD GDRIAIISHGKLCCVGSSLFLKNOLGTGYYL	1	ľ	{		í	ľ	
YSRRGIWELLLKYRQGRTIILSTHHMDEAD GDRIAIISHGKLCCVGSSLFLKNOLGTGYYL				1	l	ļ	EKHVKAEMEQMALDVGLPSSKLKSKTSQLS
GDRIAIISHGKLCCVGSSLFLKNOLGTGYYL			l		l	[GGMQRKLSVALAFVGGSKVVILDEPTAGVDP
GDRIAIISHGKLCCVGSSLFLKNQLGTGYYL	1	j	J		!	İ	YSKRGIWELLLKYRQGRTIILSTHHMDEADVL
			ŀ		ľ	l	GDRIAIISHGKLCCVGSSLFLKNQLGTGYYLT
							LVKKDVESSLSSCRNSSSTVSYLKKEDSVSQS
			ł	ļ		į	SSDAGLGSDHESDTLTIDVSAISNLIRKHVSEA
	1	1	į	1	į	j	RLVEDIGHELTYVLPYEAAKEGAFVELFHEID
						l	DRLSDLGISSYGISETTLEEIFLKVAEESGVDA
ETSDGTLPARRNRAFGDKQSCLRPFTEDD	1					l	ETSDGTLPARRNRRAFGDKQSCLRPFTEDDA
ADPNDSDIDPESRETDLLSGMDGKGSYQVK			-		İ	İ	ADPNDSDIDPESRETDLLSGMDGKGSYQVKG
	L						WKLTQQQFVALLWKRLLIARRSRKGFFAQIV

SEQ ID No.: of nucl- order bod peptide color of nucl- order by peptide colored or nucleotide unclockide uncloc				Lanc	1 5 11 1		
nucleotide sequence uneed when the contraction of t					1	I .	
Sode Seq uence	1 1		noa	1			
uence 09496 correspond 1							
uence 914 ng to first amino acid residue of peptide sequence residue of peptide sequence peptide sequence residue of peptide sequence peptide sequence residue of peptide sequence peptide sequence residue of peptide sequence peptide sequenc	1 1			1			
minto acid cesidue of peptide sequence propried sequence purposite successions. V-Usuline, W-Stop codon, /-possible nucleotide deletion, *-possible nucleotide	1 - 1	uence	1	1		l .	
residue of peptide sequence	uence		ļ	914			
peptide sequence Poposible nucleotide deletion, Poposible nucleotide insertion Poposible nucleotide insertion Poposible nucleotide insertion Poposible nucleotide insertion Poposible nucleotide insertion Poposible nucleotide insertion Poposible nucleotide insertion Poposible nucleotide insertion Poposible nucleotide insertion Poposible nucleotide insertion Poposible nucleotide Poposible	1 1						
Incidentide insertion	1					sequence	
LPAYPYCIALYPSILYPPFGKYPSILELQPWMY NGOYTFVSNDAPEDT GILELINALTEDOFG TRCMEGNPIPDTPCQAGEEWTTAPVPQTIM DLFQMGNWTMONPSACQCSSRKKKMLPV CPPGAGGLPPPGRKQNTADLIOLITGRNIBDY LVKTYVQIIAKSLKNKIWNEFRYGGFSLOVS NTQALPPSQEVMDATKQMKRKHLKAKDSSA DRFLNSLGREMTGLDTIRNINVXWFNKOW HAISSPLAWINNALRANLQKGENFSHYGGTA MSPVPASFVYFLIQREVSKAKHLPJFSGVFWI YWLSNFVWDMCNYVVPATLVIIIFICFQQKSY VSYSTIN_PVLALLLLLYGWSITPLMYPASFVFK IPSTAYVVLTSVNLFIGMGSVATFVLELFTIDN KLNINDILKSVFLIPPFICLGGCLIDWKNQ AMADALERFGENFFYSLSWULVGRNLFAM AVEGVYFLITVLIQYFTFICAFGCLIDWKNQ AMADALERFGENFFYSLSWULVGRNLFAM AVEGVYFLITVLIQYFTFIRRPRVAKLSPLN DEDEDVRREQRILDGGGQNDILEIKELTKIY DEDEDVRREQRILDGGGQNDILEIKELTKIY RKRRFAVDRICVGIPFGGFCELLGWKNG SSTFKMLTGDTTVTRGDAFLNRSILSNIHEV HQNMGVCPOPDATTELLTGREVFFFALLG VYEKEVGKVGEWARKLGLVKYGEKYAGNY SGGNKRLSTAMALIGGPPVVLDEPTTOMD PFARRFLWNCALSVVLTSSMERE ENTENDAMENT ENT	1						
NEOYTFVSNDAPEDIGTILELINALTXDPGFG TRCMGNPIPIPIPOAGEEWITANTXDPGFIG DILFONGNWTMONPSPACOCSSDKIKKMLPV CPPGAGGLPPORKONTADILODITGRISDY LVKTYVQILAKSLKNIKWNEERYGGFSLGVS NTQALPPSQEVNDATKQMKKHLKLAKDSSA DIFENSLGREMTGLDTIRNINVKWFINNKOW HAISSILNVINNALIRANLQKGENPSHYGITAF MIPINILIKQQLSEVAPHTISTVDLVSICVIFA MSFVPASFVVTLQERVSKAKHLQFISGVKPVI YWLSNFVWDMCMYVYATI.VIIIPCFQKSY VSSTNLPVLALLLLYGWSITPLMYPASFVFK PISTAYVVLTSVNLFIGINGSVATFVULFLFTDN KLNNINDILKSVFLIPPHFCLGRGLIDMYKNQ AMADALEFGENBFYSPLSWDLVGRINFAM AVEGVYFFLITVLIQVRFIRERPVNAKLSPLN DEDEDVREREQRILDGGGONDLIERIELIKTIY RRKRFAAVDRICVGIPPGECFGLLGVNGAGK SSTFKMLTGDTYTVRGDAFINNSLISNIFEV HQNMGYCPQFDAITELLTGREHVEFFALLRG VPEKEVGKVGEWARKLGLVKYGEKYAGNY SGNKRKLSTAMALIGGPPVVFLDEPTTOMD PFARRFLWNCALSVVKEGRSVVLTSHSMEEC EALCTRAMIMYNGRFCLGSVQHLKNNRFGD GYTIVVRIAGSNPDLKPVQDPFGLAFPGSVPR BEHRNMLQVQLPSSLSLARFISLSQSKKRLH IEDYSVSQTTLDQVFVNFAKDQSDDDHLKDL SHKNQTVDDVALTSFLQDEKVKESTVY TILBGVSIIGTIGAGIFISPKGVLQNTGSVGM STTWTVCOVLSLEGALSVALGTFINSGSKKRLH IEDYSVSQTTLDQVFVNFAKDQSDDDHLKDL LAFGRYTLEPFFQCEBLAKLTLAVGTTVM VLNSMSVSWSARIQIFLTFCKLTALITIPFGV MQLIKQTQNFKDAFSGRNSSTRLPLAFYYG MYAYAGWFYLNFYTERJCBEKVKGTVW VLNSMSVSWSARIQIFLTFCKTATILITIPFGV MQLIKQTQNFKDAFSGRNSSTRLPLAFYYG MYAYAGWFYLNFYTENASCFGSMNGGVFAV SRLLGNSLAYPHASTGFGSNKGGVFAV SRLLGNSLAYPHASTGFGSNKGGVFAV SRLLGNSLAYPHASCFGSNKGGVFAV SRLLGNSLAYPHASCFGSNKGGVFAV SRLLGNSLAYPHASCFGSNKGGVFAV SRLLGNSLAYPHASCFGSNKGGVFAV SRLLGNSLAYPHASCFGSNKGGVFAV SRLLGNSLAYPHASCFGSNKGGVFAV SRLLGNSLAYPHASCFGSNKGGVFAV SRLLGNSLAYPHASCFGGNKGGVFAV SRLLGNSLAYPHASCFGSNKGGVFAV SRLLGNSLAYPHASCFGSNKGGVFAV SRLLGNSLAYPHASCFGSNKGGVFAV SRLLGNSLAYPHORASCFGORNGGOPFA SRLLGNSLAYPHORASCFGORNGGOPFAC LAPRGGGLLSSRLSAGKPPRTSFGSWGVLPP MOKKPRWFRIMSEKITRILQILLEVYPEEDKL LAPRGGGLLSRLSAGGFORTSFCTENGENCL CROPNSELLRAGGHKASTATEDGENCU CROPNSELLRAGGHKASTATEDGENCU CROPNSELLRAGHKKARKVGIGKQAPDGGLGLGTVN VRDKKPATLIDPVSQDALPRKQNPQTLQLLIK KKSLAGAGGLLKSSVTENGENCU CROPNSELRAGGHKASTATEDGENCU CROPNSELRAGGHKSSTATEDGENCU CROPNSELRAGGHKSSTATEDGENCU CROPNSELRAGGHKGATENSTATENGENCU CROPNSELRAGGHKTLSTATENGENCU CROPNSELRAGGHKTLSTATENGENCU CROPN			 		sequence		
TRÉMEGNPIPDTE COAGEER THAP PYPOTIVE DE PRONONTMONPS ACQUESTICA PROPERTIES PRO	i i			l			
DLFQNGNWTMQNPSPACQCSSDKIKKMLPV CPGAGGLPPPGKONTADILQTRINISDY LVXTYYQIIAKSLKNKIWVEFRYGGFSLQ NTQALPPSQEVINDATLQMKKHLAKADSSA DRFLNSLGRFMTGLDTRINVKVWFNNKGW HAISSFLNVDNALLRAILQKGENPSHYGTIFA HHLAILTKQQLSVAPMTTSVDVLVSICVIFA MSFVPASFVVFLIQERVSKAKHLQFISGVKVDV YWLSINVWDMCNVVPATLUHIFICFQQKSY VSSTRLFVLALLLLLYGWSTFLMYPASFVFK IPSTAYVVLTSVNLFIGINGSVAFTLELFTDN KLNNINDILKSVFLIPHFECGRGLDMYKNO AMADADALERFGENFVSPLSWDLVGRILFAM AVGGVVFFLITVLIQYKFFIRPRFVNAKLSPIL DEDEDVRERGRILDGGQNDDLIKBLTKIV RRKKPAVDRICVGIPPGECFGLLGVNGAGK SSTFKFMLTGDTTVTRQDAFLNKNSILSNHFEV HQNMGYCPQFDATTELLTGREHVEFFALLRG VPEKEVGKVGEWARKLGLVKYGEKVAGNY SGGNKRKLSTAMALIGGPPVVFLDEFTTGMD PFARRFLWNCALSVVEEGRSVVLTSHISMEE EALCTRAMINWORFRCLGSVQHLKNRFGD GYTIVVRIAGSNPDLKPVQDFFGLAFFGSVYF EKHRMALGYQLPSSLSSLARIFSLSGSKKRLH IEDYSVSQTTLDQVFVNFAKDQSDDDHLKDL SLHKNOTVDVAVLTSFLQDEKVESVV STRUMVLTSFLQDEKVESVV STRUMVLTSFLQDEKVESVV VLNSMSVSRARIOFILTFCKLTAILHIVPGV MQLKGQTQNFKDAFSGRVQLKRKV TLLRGVSHIGTHGAGIFISFKGVLQNTGSVGK SLTWTVCGVLSFGASTAGLTNIKSGGH YTYLEVFGLPAFVRVWVELLHRPAATAVIS LAFGRYLEPFFFGCFFELAKLTHGTVFW MULKGQTQNFKDAFSGRSSTTRLPLAFFYY MYAYAGWFYLNFVTEEVEPFEKTPLAICISM AVTIGVYLNVAYFTTINAEELLLSNAVAV SERLIGNFSLAVPIVALSFLOREKFTQGEKYQLKKV TLRMJSHIGTHGAGIFISFKGVLQNTGSVGV NGLIKGQTQNFKDAFSGRSSTTRLPLAFFYY MYAYAGWFYLNFVTEEVEPFEKTPLAICISM AVTIGVYLNVAYFTTINAEELLLSNAVAV SERLIGNFSLAVPIVALSFLOREKTPQGEKVQLKRV VLNSMSVSNARIOFILTFCKLTAILHIVPGV MQLKGQTQNFKDAFSGRSSTTRLPLAFFYY MYAYAGWFYLNFVTEEVEPFEKTPLAICISM AVTIGVYLNVAYFTTINAEELLLSNAVAV SERLIGNFSLAVPIVALSFLOREKTPQGEKVQLKRV VLNSMSVSNARIOFILTFCKLTAILHIVPGV MQLKGQTQNFKDAFSGRSSTTRLPLAFFYY MYAYAGWFYLNFTEEVEPFEKTPLAICISM AVTIGVYLNVAYFTTINAEELLLSNAVAV SERLIGNFSLAVPIVALSFLOREKTPQGEVQLF LADAASMSGVARVARISIESACKEQEVUFGDL AVAGLYLRYKCPDMHRPRKVPLTPALFSFT LEMVALSLYSDFTGGSVTTTCTGPAYYTLFI WDKKPPMTPRASPFGGSWOTKLEAGER EFAAGTGGAOQEWFGGSSADODEGGVKL REGGGLLSSRLSAGKPPLTSFTGGSWGVLYD- LADAASMSGVARVSIESACKEQEVKLFGLSLX KALAGAQALLKAGARLTKSVTENQEKKL REGGGLLSSRLSAGKPPLTSFTGGSWGVLYD- LADAASMSGVARVARISIESACKEQHEVGLD GTETYLPPLSMSGONLARLAGRDFSGGSGSGE EFAAGTEGAOQEWFGAGSSADODEGGVKL REFERVERVENGENGTHT							1 3
CPPGAGGLPPQRKQNTADIQLICTGRNISDY LVKTYYQIIASKIKNIWVEFRYGGFSLGVS NTQALPPSQEVNDATKQMKKHLKLAKDSSA DRFLNSLGRENTGDTRNNVKYPHNKGW HAISSFLNVDNAILRANLQKGENFSHYGITAF NHFLNLTKQQLSEVAPMTTSVDVLVSICVIPP AMSVEVASFVYTLGERVSKAKHQPISGVKPVI YWLSNEVWDMCNYVVPATLVIIIFICFQQKSYVPVI YWLSNEVWDMCNYVVPATLVIIIFICFQQKSYVPVI YWLSNEVWDMCNYVVPATLVIIIFICFQGVSVKPVI YWLSNEVWDMCNYVVPATLVIIIFICFQGVSVKPVI KLNNINDILKSVFLIPPHFCIGGGIDDMYKNQ AMADALERGENFVSPLSWDLVGRNLPAM AVEGVVFFLITVLQVFFIERPFVAKLSPLN DEDEDVRRERQRILDGGQONDILEIKELTKIY REKRKPAVDRICVGVPREFFALLGE VPEKEVEKVGEWARKLGLVKYGEKYAGNY SGGNKRKLSTAMALIGGPVVFLDEFTTGMD PRARRFLWNCALSVVKEGRSYVVLSTISMEC EALCTRMAIMVNGFFRCLOSVQHLKNRFGD GYPTIVVRIAGSNPDLFVQDFFCLAFFGSVPK EKHRNMLQYQLPSSLSSLARIFSILSQSKKRL IEDVSVSQTTLDQVFVNFAKDQSDDPHLKDL SLHKNQTVVDVAVLTSFLQDEKVKESVV TLRGVSINGTTIGAGSTRIFT STEPPEVFTFWARFVSTSKGG YLQGNVNGRLPSLOPEVFTFVTFVTFVTSKGG YLQGNVNGRLPSLOPEVFTFVTFVTSKGG YLQGNVNGRLPSLOPEVFTFVTFVTSKGG YLQGNVNGRLPSLOPEVFTFVTFVTSKGG YLQGNVNGRLPSLOPEVFTFVTFVTSKGG YLQGNVNGRLPSLOPEVFTFVTFVTSKGG YLQGNVNGRLPSLOPEVFTFVTFVTSKGG YLQGNVNGRLPSLOPEVFTFVTFVTSKGG YLQGNVNGRLPSLOPEVFTFVTFVTFVTSKGG YLQGNVNGRLPSLOPEVFTFVTFVTSKGG YLQGNVNGRLPSLOPEVFTFVTFVTFVTSKGG YLQGNVNGRLPSLOPEVFTFVTFVTSKGG YLQGNVNGRLPSLOPEVFTFVTFVTSKGG YLQGNVNGRLPSLOPEVFTFVTFVTSKGG YLQGNVNGRLPSLOPEVFTFVTFVTSKGG YLQGNVNGRLPSLOPEVFTFVTFVTSKGG YLQGNVNGRLPSLOPEVFTFVTFVTSKGG YLQGNVNGRLPSLOPEVFTFVTFVTSKGG YLQGNVNGRLPSLOPEVFTFVTFVTSKGG YLQGNVNGRLPSLOPEVFTFVTFVTSKGG YLQGNVNGRLPSLOPEVFTFVTSKGG YLQGNVNGRLPSLOPEFTTFLACISM AVTILEVFGPLFFTTLACISM AVTILEVFGPLFFTTLACISM AVTILEVFGPLFFTTLACISM AVTILEVFGPLFFTTLACISM AVTILEVFGPLFFTLACISM AVTILEVFGPLFFTTLACISM AVTILEVFGPLFFTLACISM AVTILEVFGPLFFTLACISM AVTILEVFGPLFFTLACISM AVTILEVFGPLFFTLACISM AVTILEVFGPLFFTLACISM AVTILEVFGPLFFTLACISM AVTILEVFGPLFFTLACISM AVTILEVFGPLFFTLACISM AVTILEVFGPLFTLACISM AVTILEVFGPLFTLACISM AVTILEVFGPLFTLACISM AVTILEVFGPLFTLACISM AVTILEVFGPLFTLACISM AVTILEVFGPLFTLACISM AVTILEVFGPLFTLACISM AVTILEVFGPLFTLACISM AVTILEVFGPLFTLACISM AVTILEVFGPLFTLACISM AVTILEVFGPLFTLACISM AVTILEVFGPLFTLACISM AVTILEVFGPLFTLACISM AVTILE							
LVKTYVOJIAKSLKNIKIWVNEFRYGGFGLORNOK WITOALPPSGEVINDATKOMKKHLKLAKDSSA DRFLINSLGREMTGLOTENNVKWENNK GEN HAISSELNVENDALKANLERAL GREGORENSTYGTAF MEPLALTKOQLSEVAPMTISVDVLVSIGVIFA MSFLVPASFVVPLIQER YSKAKHLQFISGVKPVI YWLSNEVWDMCNYVVPAILUHIGEPQGKSY VSSTNLEPVLALLLLLYGWSITELMYPASFVYF IPSTAYVVLTSVNLFGINGSVATFVLELFITON KLINNINDILKSVFLEPHFELGRGLDMYKNON AMADALERTGERRYSPISSUDLVGRILFAM AVEGVYFFLITVLIQYRFFERPRYVNAKLSPIL DEDEDVRERGRILDGGGONDLIKELTKIY RRKKRPAVDRICVGIPPGECFGLLGVNOAGK STFKMLTGDTTYTRGDAFLNSRNLISNIHEV HQNMGYCPOPDAITELLTGREHVFFFALRG VPEKEYGKVGEWARKLGLUKVGEKYAGNY SGGNKRKLSTAMALIGGPVVFLDEFTTOMD PKARRFLWVOJESUSSVNLTSSMEEC EALCTRAMINVORFRCLGSVOHLKNRFGD GYTIVVRIAGSSPPLLKVPOPFGLAFFGSWYF EKHRNMLQYQLPSSLSSLARIFSLSQKKKLH IEDYSSVGTTLDQVFVNFAKDQSDDDHLKDL SHKNOTVODVAUTSFLQDEKVESTVV THE STREET			İ				
PREVINSIEGE PROPERTIES OF PREVINCENT STORY OF THE PROPERTIES OF PREVINCENT STORY OF THE PROPERTIES OF PROPERTIES O]		ļ]			
DRFLNSLGRFMTGLDTRNNVKVWFNNKGW HAISSELNVINALIRANLQSENPSHYGITA MSFVPASFVVLIQERVSKAKHLQFISGVKPW WSSNPVASFVVLIQERVSKAKHLQFISGVKPW YWLSNFVWDMCNYVVPATLVIHFICFQQKSY VSSTNLPVLALLLLYGWSITPLMFYASFVFK IPSTAYVVLTSVNLFIGINGSVATFVASFVFK IPSTAYVVLTSVNLFIGINGSVMTPVASFVFK IPSTAYVVLTSVNLFIGINGSVMTPVASFVFK IPSTAYVVLTSVNLFIGINGSVMTPVASFVFK IPSTAYVVLTSVNLFIGINGSVMTPVASFVFK IPSTAYVVLTSVNLFIGINGSVMTVNQ AMADALERFGENFVSPILSVDLVGRNLFAN AVEGVVFFLITVLIQVRFFRPRPVNAKLSPLN DEDEDVRRERGRILDGGGDNLEIKELTKIV RRKRKPAVDRICVGIPPGECFGLLGVNGAGK SSTFKMLTGDTTVTRGDAFLNRNSILSNIHEV HQNMGYCPGPAITELLTGREHVEFFALLRG VPEKEVGKVGEWAIRKLGLVKYGEKYAGNY SGGNKRKLSTAMALIGGPPVFLDEPTTGMD PKARRFLWNCALSVVKEGRSVVLTSHSMEC EALCTRMAIMVGNFRCLGSVHLKNFGGD GYTTVVRIAGSNPDLKPVQDFFGLAFFGSVPK EKHRNMLQVQLPSSLSSLARISJGSKKKH IEDYSVSQTTLDQVFVVFAKADQSDDDHKDL SLHKNQTVVDVAVLTSFLQDEKVKESYV 1977 2327 A \$260 3 1567 IPGSTISFSLCFIFPPCVFTMVKRPVVSTISKGG YLQGNVNGCLPSLGSAKKESYV TLLEGVSIIGFTIGAGFIFISKGVLQNTGSVGM SLTMTVCGVLSLFGALSVAELGTHKKSGGH YTVILEVFGPLAFFVRVWELLIRRAATAVIS LAFGRYILEPFFIQCEIPELAIKLITAVGITVVM VLNSMSVSWSARQIQFLTFCKLAILIIVPCV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTTPLAICISM AIVTIGTVLTINVAYFTTINAEBLLISNAVAVT PSERLLGNFSLAVPIFVALSCGGSMGGYFAV SRLFYVASREGHLPELLSMIHVRKHTPLPAVIV LHPLTMIMIFSGDLDSLLNFLSFARWLFGIA VAGIYLRYKCPDMHRPREVYPALFICIA VAGIYLRYKCPDMHRPREVYPALFICIA VAGIYLRYKCPDMHRPREVPAYLFIF UDKKPRWFRIMSKITTRLQUEVPEEDALL LAFGRYLLARIMSKITTRLQUEVPEEDAL LAFGRYLLARIMSKITTRLQUEVPEEDAL LAFGRYLLARIMSKITTRLQUEVPEEDAL LAFGRYLLARIMSKITTRLQUEVPEEDAL LAFGRYLLARIMSKITTRLQUEVPEEDAL LAFGRYLLARIMSKITTRLQUEVPEEDAL VAGIYLRYKCPDMHRPREVPRONPOTLQLISK KKSLAGAGOLLKGAERLTKSVTENGENKLO ROPNSELLRAQHWKRKYGDRADJELGVRN AGSLPPHIGTTEVINNTDLLDLKKIPEDVCPI. VRIKKFMTLDPVSQDALPPRONPOTLQLISK KKSLAGAAGOLLKGAERLTKSVTENGENKLO ROPNSELLRAQHWKRKYGDRADVHLCKEG LAFRAGTEDADAWNLCKEG LAFRAGTEDADAWNLCKEG LAFRAGTEDADAWNLCKEG LAFRAGTEDADAWNLLCKEG LAFRAGTEDADAWNLCKEG LAFRAGTEDADAWNLCKEG LAFRAGTEDADAWNLCKEG LAFRAGTEDADAWNLCKEG LAFRAGTEDADAWNLCKEG LAFRAGTEDADAWNLCKEG LAFRAGTEDADAWNLCKEG LAFRAGTEDADAWNLCKEG LAFRAGTEDADAWNLCKEG LAFR							
HAISSENVINNALRANLOGENPSHYGTIAF MHPLNITKO, QLSEVAPMTISVDVLVSICUIFA MSFVPASFVVFLIQERVSKAKHLQFISGVKPVI YWLSNFVWDMCNYVVPATLVIIHIECOGKSY VSSTNLPVLALLLLUS WSSTPLMYPASFVFK IPSTAYVVLTSVNLFIGINGSVATFVLLEITDN KLNNINDILKSVFLIPFHFECGEGLIDMYKNO, AMADALERFGENFVSPLSWDLVGRNI,FAM AVEGVVFFLITVLIQVFRETFRYPVAKLSPIN DEDEDVRRERQRILDGGGQNDILEIKELTKIV RRRKRFAVDRICVGIPPGECFGLLGVNGAGK SSTFKMLIGDTIVTRGDAFINSLISNIHEV HQNMGYCPQFDAITELLTGREHVEFFALLG VPEKEVGKVGEWAIRKIGLIKYYGEKYAGNY SGGNKRKLSTAMALIGGPVVFLDEFTIGMD PKARRFLWNGALSVVKEGREV GYTIVVRIAGSNIPLKYQGFFGLAFFGSVPK EKHRNMIQYQLPSSLSSLARIFSILSGSKKRLI EDYSVSQTITLDQVFVVFLAFSTSFGG GYTIVVRIAGSNIPLKPVQDFFGLAFFGSVPK EKHRNMIQYQLPSSLSSLARIFSILSGSKKRLI EDYSVSQTITLDQVFVNFAKDQSDDDHIKDL SLIKNQTVVDVAVLTSFLQDEKVKESYV TLLRGVSIIGTIGAGFISPKGVLQNTGSVGM SLTIMVTVGVLSLEFGALSYAELGTITLKKGGH YTTVLEVFGPLPAFVRVWVELLIIRRAATAVIS LAFGRYILEPFFIQCEIPELAIKLITAVGITVVM VINSMSVSWARRQIGFLITEKLALILIIVPCV MQLIKGQTQNFKDAFSGRSSITRLPLAFYG MYAYAGWFYLNFVEEVENPEKTIPLAFYGM MYAYAGWFYLNFVTEEVENPEKTIPLAFYGM MYAYAGWFYLNFVTEEVENPEKTIPLAFYGH MYAYAGWFYLNFVTEEVENPEKTIPLAFYGH MYAYAGWFYLNFVTEEVENPEKTIPLAFYGH WKRPRWFRIMBSKITRTLQIILEVVYEEDKAL VAGLIYLRYKCPDMHRPRKYATYLFIA VAGLIYLRYKCPDMHRPRKYATYLFIA VAGLIYLRYKCPDMHRPRKYDFILLSFGRGWPLPFGLL RHYGGGLLSSRLSAGKPPLRTSFFGSWOVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FOPSLWPWDSVRNNLRSALTEMCULVDVLSI VRUKKFMTLDPVSQDALPPKONPOTLOLISK KKSLAGAAQILLKGAERLTKSVTENGENKLQ RDFNSELLRLRQHWKLRKVDGNLIGDLSVYRS AGSLPPHIGTTEVINNTDLLDLKKIPEDVCPI. DVQIPSOLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLKSKRPSPHYOTILLGLDLKNIPEDVCPI. DVQIPSOLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLYKSRPSPHYOTILLGLDLKNIPEDVCPI. DVQIPSOLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLKSKRPSPHYOTILDLIGKRIPEDVCPI. DVQIPSOLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLKSKRPSPHYOTILDLIGKNIPEDVCPI. DVQIPSOLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLKSKRPSSPHYOTILGTGRANVLLCKEI LFKRPLKSKRPSSPHYOTILGTGRANVLLCKEI LFKRPLKSKRPSSPHYOTILGTGRANVLLCKEI LFKRPLKSKRPSSPHYOTILTGRANVLLCKEI LFKRPLKSKRPSSPHYOTILGTGRANVLLCKEI LFKRPLKSKRPSSPHYOTILTGRANVLLCKEI LFKRPLKSKRPSSPHYOTILTGRANVLLC							
NHPLNLTKQQLSEVAPMTSVDVLVSIVIFICA							1
MSFVPASFVVĪLIGERVSKAKHLOFISGVKPYV YWLSNFVVMMCNYV VEWDMCNYV JUIFICOQKSY VSSTNLPVLALLILLYGWSITPLMYPASFVFK IPSTA AVVLTSVNLPIGINGSVATFVLEIFTIDN KLINNINDILKSVFLIPHELGRGLIDMYKNQ AMADALERGENRFVSFLSWILVGRNLFAM AVEGVVFFLITVLIQVRFIRPRPVNAKLSPLN DEDEDVRERGRILDGGGQNDILEIKELTKIY RRKRFPAVDRICVGIPPEGFLLGWNGAKCK SSTFKMLTGDTTVTRGDAFLNRNSILSNIHEV HQNMGYCPQFDAITELLTGREHVEFFALLRG VPEKEVGKVGEWARKLGVKYGEKYAGNY SGGNKRKLSTAMALIGGPPVVFLDEPTIGMD PKARRFL WNCALSVVKGENVYSISMEC EALCTRMAIMYNGFRCLGSVOPH.KNRFGD GYTIVVELGSNPDLKEYQDFFGLAPFGSVPK EKHRNMLQYQLPSSLSSLARIFSILSQSKKRLH IEDYSVSQTTILDQVFVNFAKDQSDDDHLKDL IEDYSVSQTTILDQVFVNFAKDQSDDDHLKDL SHKNQTVVDVAVLTSFLQDEKVKESYV 1LRGVSINGTHIGAGIFSPKGVLQNTGSVGM SLIHKNQTVVDVAVLTSFLQDEKVKESYV TLRGVSINGTHGAGIFSPKGVLQNTGSVGM SLITHTVCQTSLSFGALSYAELGTTIKKSGGH YTYILEVFGPLPAFVRVWVELLIRPAATAVIS LAFGRYILEPFFIQCSIPELAIKLITAAGTIVVM VLNSMSVSWSARQIGHLFTCKLTALILIIVPGV MQLIKGQTQNFKDAFSGRDSSITELJAFYYG MYAYAGWYLNFVTESVENPEKTIPLAFYYG MYAYAGWYLNFVTESVENPEKTIPLAFYYG MYAYAGWYLNFVTESVENPEKTIPLAFYYG MYAYAGWYLNFVTESVENPEKTIPLAFYYG MYAYAGWYLNFVTESVENPEKTIPLAFYYG MYAYAGWYLNFVTESVENPEKTIPLAFYYG MYAYAGWYLNFVTESVENPEKTIPLAFYGH AVAGLIYLRYKCEDMHRFWLFVLFPALFSFTC LFMVALSLYSDPFSTGIGFVTTLUILGVPYEEDKL WDKKPRWFRIMSEELITRSFSTGSWGVLPP FSERLIGNFXXCEDMHRFWKHTPLIFFAFSTC LFMVALSLYSDPFSTGIGFVTTLUILEVPPEDKL MOKKPRWFRIMSEELITRSFTGSWGVLPP LADAASMSGVRAVRISESACEKQVHEVGLL GTETYLPPLSMSQNLARLAQRIDFSGGSSEE EEAAGTTEGDAQEWFGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPRQNPOTIQLISI VRDKKFMTLDPVSQDALPPRQNPOTIQLISI VRDKKFMTLDPVSQDALPPRQNPOTIQLISI VRDKKFMTLDPVSQDALPPRQNPOTIQLISI KKSLAGAAQILLKGAERLTKSVTENQENKLQ ROFNSELLRLRQHWKLRKVGDKLGGLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDVCPL DVQIPSDLEGSAYIKVSIQKQAPDIDGLGTVN FLYRPLFSKFGSSHYGNLEAAGNVLLCKEI							· · · · · · · · · · · · · · · · · · ·
YWLSNFVWDMCNYVPATLVIIIFICFQQESYY VSSTIN_PVALAILLI YGWSITPIMPYASFYV IPSTAYVVLTSVNLFIGINGSVATFVLELFIDN KLNNINDILKSVFLIFPHFCLGRGLIDMYKNQ AMADALERFGENRFVS-BLSWDLVGRNIFAM AVEGVVFFLITVLIQYRFFIRPPVAKLSFLN DEDEDVRREQRILDGGQONDILEIKELTKIY RKKRFAVDRICVGIPPGECFGLLGVNGAGK SSTFKMLTGDTTVTRGDAELNINSILSNIHEV HQNMGYCPQFDAITELLTGREHVEFFALLRG VPEKEVGKVGEWAIRKLGLVKYGEKYAGNY SGGNKRKLSTAMALIGGPVYCEGKYAGNY SGGNKRKLSTAMALIGPVYCEGKYAGNY SGGNKRKLSTAMALIGPVYCEGKYAGNY SGGNKRKLSTAMALIGPVYCEGKYAGNY SGGNKRKLSTAMALIGPVYCEGKYAGNY SGGNKRKLSTAMALIGPVYCEGKYAGNY SGGNKRKLSTAMALIGPVYCHOPYLDEPTTGMD PKARRFLWCALSVVKEGRSVVLTSHSMEEC EALCTRMAMYNGRFRCLGSVQHLKNRFGI GYTIVVRIAGSNPDLKPVQDFFGLAFPGSVPK EKHRNMLQYQLPSSLSSLARIFSILSGKKRLH IEDYSVSQTTILDQVFVNFAKDQSDDHLKDL SLHKNQTVVDVAVLTSFLQDEKVKESYV PSSTISFELTGPPVVPTWAKPVVSTISKGG YLQGNVNGRLPSLGNKEPPQGEKVQLKRV TLRGYSIIIGTIIGAGIFSPRGVLQDNTGSVCM SLITWTVCQVSLSFGALSYAELGTTIKKSGGH YTYILEVFGPLPAFVRVWVELLIRPAATAVIS LAFGRYLIEPFFQCEFELSHALTIAVGTVVM VLNSMSVSWSARIQIFLTFCKLTAILIIVPCV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTESVENPEKTIFTOTMOTMOTMOTMOTMOTMOTMOTMOTMOTMOTMOTMOTM	1 1			l			
VSSTIN_PVLALLLLIYGWSITPLMPPASFVFK IPSTAYVVLTSVMLFIGINGSVATFVLELFIDN KLNNINDILKSVFLIFPHFCLGRGLIDMVKNQ AMADALERGENRFVSFLSWDLJVGRNLFAM AVEGVVFFLITVLIQYRFIRPRPVNAKLSPLN DEDEDVRERGRILDGGGQNDILEIKELTKIY RRKRFAPAVDRICVGIPPGEFGLLGVWGAGK SSTFKMLTGDTTVTRODAFLNRNSILSNIHEV HQNMGYCPQFDAITELLTGREHVEFFALLRG VPEKEVGKVGEWAIRKLGVKYGEKYAGNY SGGNKRKLSTAMALIGGPPVVFLDEPTTGMD PKARRFL WICA LSVVKGESVVLTSHSMEC EALCTRMAIMVNGRFRCLGSVQHLKNRFGD GYTIVVRIAGSNPDLKPVQDFFGLAFGSVPK EKHRNMLQYQLPSSLSSASVVLTSHSMEEC EALCTRMAIMVNGRFRCLGSVQHLKNRFGD GYTIVVRIAGSNPDLKPVQDFFGLAFFGSVPK EKHRNMLQYQLPSSLSSLARIFSILSQSKKRLH IEDYSVSQTTLDQVFVNFAKDQSDDHLKDL SLHKNQTVDVDVAVLTSHQDEKVKESYV TLRGVSIIIGTTIGAGFFFGVPFTMVRKFVVSTTSKGG YLQGNVNGRLPSLGNKEPPGQEKVLKRKV TLRGVSIIIGTTIGAGFFFGVPTMVRKFVVSTTSKGG YLQGNVNGRLPSLGNKEPPGQEKVLKRKV TLRGVSIIIGTTIGAGFFFGVPTMVRKFVVSTTSKGG YLQGNVNGRLPSLGNKEPPGQEKVLKRKV TLRGVSIIIGTTIGAGFFFGVPTMVRKFVVSTTSKGG YLQGNVNGLPSLGNKEPPGQEKVLKRKV TLRGVSIIIGTTIGAGFFTGVFTTNAFELLISNAVAVT PSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPELSMHVRKHTPLAVIV LHPLTMMLFSGDLDSLLNFLSFARWLFFGLA VAGLIVLRVKCPDMHRPFKVPLFPLAFSTC LFMVALSLYSDPPSTGIGFFTTLPLAFSTC LFMVALSLYSDPSTGIGFFTTLPLAFSTT LFMVALSLYSDPSTGIGFFTTLPLAFSTT LFMVALSLYSDPSTGIGFFTTLPLAFSTT LFMVALSLYSDPSTGIGFFTTLPLAFSTT LFMVALSLSTRLSRAKFTRLQHLEVVPEEDKL RGGGLISSRLSAKFPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKGVPEVGLD GTETYLPPLSMSQNLARLAQRIDPSGGSGSE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNILRSALTEMCVLTVDLLSI VRDKKFMTLDPVSQDALAPPKQNPQTILQUISK KKSLAGAAQILLKGAERLTKSVTTSQENKLQ RDFNSBLLBRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSTGKRAGNPOTLQUISK KKSLAGAAQILLKGAERLTKSVTTSQENKLQ RDFNSBLLBRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSTGKRAGNPOTLQUISK]			l			l
PSTAYVVLTSVNLPIGINGSVATFVLELFIDN KLNNINDILKSVFLIPHPICLGRGLIDMYKNQ AMADALERGENRFVSPLSWDLVGRNLFAM AVEGVYFFLITVLIQYREFIRRPRVNAKLSPLIN DEDEDVRERQRILDOGGONDILEIKELTKIY RKKKPAVDRICVGIPPGECFGLLGVNGAGK SSTFKMLTGDTTVTRGDAFLNRNSILSNIHEV HQNMGYCPQFDAITELLTGREHVEFFALLRG VPEKEVGKVGEWARKLGJLVK YGEKYAGNY SGGNKRKLSTAMALIGGPPVVFLDEPTTGMD PKARRFLWNCALSVVKEGRSVVLTSHSMEEC EALCTRMAIMYNGFRECLGSVQHLKNRFGD GYTIVVRIAGSNPDLKPVQDFFGLAFPGSVPK EKHRMLQYQLPSSLSSLARIFSILSQSKKRLH IEDYSVSQTTILDQVFVNFAKDQSDDDHLKDL SLHKNQTVVDVAVLTSFLODEKVKESYV PGSTIFSELGFIPPVCVTPMVKPVVSTISKGG YLQGNVNGRLPSLGNKEPPGQEKVQLKRKV TILRGVSIIIGTIGAGIFISPKGVLQNTGSVGM SLTIWTVCGVLSLFGALSYAELGTTIKKSGGH YTYLEVFGPLPAFVRVWVELLIIRPAATAVIS LAFGRYLLEPFFQCEIPELSMHVRKHTPLPAVIV VINSMSVSWSARIOJFLTFCKLTAILIILVPGV MQLIKGGTQPKFKDAFSGNSTELPLAFYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYPTTINAEELLISNAVAYT FSERLIGNFSLAVPIFVALSCFGSNNGGVFAV SRLFYVASREGHLPELSMHVRKHTPLPAVIV LHPLTMMIRSGDLDSLINSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALISFTC LFMVALSLYSDPFSTIGGFVTTLTGVPAYYLFII WDKKPRWFRIMSEKITRIQIILEVVPEEDKL GTETYLPPLSWSQNLARLAQRIDFSQGSGSE EEAAGTEGDAQEWFGAGSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLTVULSIK KKSLAGAAQILLKGAERLTKSVTENGENKLQ RDFNSELLERHOPMUKLRKVGDKILGDLSYNS AGSLFPHHGTFEVKNTDLDLDKKIPEDCYPL DVQIPSDLEGSAYIKVSIQKQAPDIDGLGTVN LFKRPLPKSKFOSHWVGNLEAAQNVLLICKEI	1						l
SILNNINDILKSVFLIFPHECLGRGLIDMVKNQ AMADALERGENRFVSPLSWDLVGRNIFAM AVEGVVFFLITVLIQVKFRIRPRFVNAKLSPLN DEDEDVRRERQRILDGGGQNDLEIKELTKIY RRKRKPAVDRICVGIPPGECFGLLGVNGAGK SSTFKMLTGDTTVTRGDAFLNKNSILSNIHEV HQNMGYCPQFDAITELLTGREHVEFFALLRG VPEKEVGKVGEWAIRKLGLVKYGEKYAGNY SGGNKRKLSTAMALIGGPPVYFLDEPTTGMD PKARRFLWNCALSVVKEGRSVVLTSHSMECC EALCTRMAIMVNGRFRCLGSVQHLKNRFGD GYTIVVRIAGSNPDLKPVQDFFGLAFPGSVPK EKHRNMLQYQLFSSLSSLARIFSILSQSKKRLH IEDVSVSQTTLDQVFVNFAKDQSDDDHLKDL SHKNOTVVDVAVLTSLDDEKVKESYV SKHRHMLQYQLFSSLSSLARIFSILSQSKKRLH IEDVSVSQTTLDQVFVNFAKDQSDDDHLKDL SHKNOTVVDVAVLTSLDGEKVQLKRV TLLRGVSIIIGTIGAGFISPRGVLQNTGSVGM SLTIWTVCGVLSLFGALSYAELGTTIKKSGGH YTQLGNVNGRLPSLQDEKVQLKRKV TLLRGVSIIIGTIGAGFISPRGVLQNTGSVGM SLTIWTVCGVLSLFGALSYAELGTTIKKSGGH YTYLLEVFGPLPAFVWELLIIRPAATAVIS LAFGRYLEPFFIQCEIPPLAIKLTAVGITVVM VLNSMSVSSARIQIFITFCKLTAILIIVPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGVLTNAVFTTINAEELLISNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLNNLSFARWLFIGLA VAGLTYLTWYCFDVHIRPFKVPLFPALFSFTC LFMVALSLYSPFFSTGFFTUTLTGVPAYTLFII WDKKPRWFIRMSKITRTLQILLEVPEEDKL FMGGGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTGTYLTPLSMSQNLALQRIDFSGSGSEE EEAAGTEGDAQEWPGAGSSADQDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLISK KKSLAGAAQILLKGAGRLTKSVTENQEKLIGDLSYRS AGSLFPHTGFFVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAXIKVSIQKGAPDIGLGTVN FRORFSKPDSTFFORNAL QRDFFSGGSATIKKVSIQKGAPDIGLGTVN LFKRPLPFKSKPGSPHWGTKLEAAQNVLLCKEI				1			l
AMADALERFGENFVSPLSWDLVGRNLFAM AVEGVVFFLITVLIQYRFFIRPRPVNAKLSPLN DDEDDVRRERQRILDGGQNDILEIKELTKIY RRKRRPAVDRICVGIPPGECFGLGVNGAGK SSTFKMLTGDTTVTRGDAFLNRNSILSNIHEV HQMMGYCPQFDAITELLTGREHVEFFALLRG VPEKEVGKVGEWAIRKLGLVKYGEKYAGNY SGGNKRKLSTAMALIGEPVVFLDEPTTGMD PKARRFLWNCALSVVKEGRSVUTSHSMEEC EALCTRMAMWNGRFRCLGSVQHLKNRFGD GYTIVVRIAGSNPDLKPVQDFFGLAFPGSVPK EKHRNMLQYQLPSSLSSLARIFSILSQSKKKLH IEDYSVSGYTILDQVFNAKDQSDDPHLKDL SLHKNQTVVDVAVLTSFLQDEKVKESYV 19327 A 8260 3 1567 IPGSTISFSLCFIFPPCVPTMVKPVVSTISKGG YLQGNVNGRLPSLGNKEPPGQEKVQLKRKV TILLRGVSIIIGTIIGAGIFISPRGVUQNTGSVGM SLTIWTVGQVLSLFGALSYAELGTTIKKSGGH YTYILEVFGPLPAFVRVWELLIIRPAATAVIS LAFGRYILEPFFIQCEIPELAIKLTAVGITVVM VINSMSVSWSARIQIFLTFCKLTALIIIVPGV MQLIKGQTQNFKDAFSGRDSSTIR.PLAFYG MYAYAGWFYLNFVTEEVENPEKTPLAICISM AIVTIGYVLTNVAYFTTNAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAA SRIFYVASREGHLPEIFTGCEPELAIKLITAGITVW LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPKVPLFPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL THE ANASMSGVRAVRISIESACEKQVHEVGLA ASMSGVRAVRISIESACEKQVHEVGLA GGGSLRCVLGKLLGGLFGSBFCVRPFGGL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLU GTETYLPIJSMSQNLARLAQRIDFSGGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDAJPPKQNPGTLQLISK KKSLAGAAQILLKGAGRLTKSVTENQENKLGDLSYRS AGSLFPHHGTEVKNTDLDLDKKIPEDVCL DVQHSDLEGGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWYGKLEAAQNVLLCKEI	Į į		ł	1	}		l
AVEGVVFFLITVLIQVRFFIRPRPVNAKLSPLN DEDEDVRRERQRILDGGQNDILEIKELTKIY RKRKPAVDRICVGIPPGECFGLLGVNGAGK SSTFKMLTGDTTVTRGDAFLNRNSILSNIHEV HQNMGVPCPDATILTGREHVEFFALLRG VPEKEVGKVGEWAIRKLGLVKYGEKYAGNY SGGNKRKLSTAMALIGGPPVVFLDEPTTGMD PKARRFLWNCALSVVKEGRSVVLTSHSMECC EALCTRMAIMVNGRFRCLGSVQHLKNRFGD GYTIVVRIAGSNPDLKYOPFGLAFPGSVPK EKHRNMLQYQLPSSLSSLARIFSILSQSKKRLH IEDYSNSQTTIDQVFVNFAKDQSDDDHLKDI- SHKNQTVVDVAVLTSFLQDEKVKESYV 977 2327 A 8260 3 1567 IPGSTISFSLCFIPPCVPTMVRKPVSTISKGG YLQGNNORGLPSLGNKEPPGQEKVQLKREV TLLRGVSIIIGTIIGAGIFISPKGVLQNTGSVGM SLITIWTVCGVLSLFGALSYAELGTTIKKSGGH YTYLLEVFGPLPAFVRVWVELLIIRPAATAVIS LAFGRYILEPFFIQCEIPELAIKLITAVGITVVM VLNSMSVSWSARJCHFCKLTALLIIVPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIFLAICISM AIVTIGYVLTNVAYFTITNAEELLISNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFVVASREGHLPEITTICKLTALLIIVPGV VAGLIYLRYKCPDMHRPRKVPLFIPALFSFTC LFMVALSLYSPPSTGIGFVTILTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL VAGLIYLRYKCPDMHRPRKVPLFIPALFSFTC LFMVALSLYSPDFSTGIGFVTILTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL HRGCGLLSSRLSAGKPLRTSSFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQOILALQRIDFSGSGSGE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNILRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAGRLTKSVTENQENLIGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDVCI. DVQHSDLEGSAVIKVSIQKAGPDISJYRS AGSLFPHGTFEVIKNTDLDLDKKIPEDVCI. DVQHSDLEGSAVIKVSIQKAGPDIGSJGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI				1			1
DEDEDVRRERGRILDGGGONDILEIKELTKIY RRKKRAVDRICVGIPPGECFGLLGVNGAGK SSTFKMLTGDTTYTGDAFLNRNSILSNIHEV HQMMGYCPGDAITELLTGREHVEFFALLRG VPEKEVGKVGEWKGEWKLGLVKYGEKYAGNY SGGNKRKLSTAMALIGGPPVVFLDEPTTGMD PKARRFLWNCALSVVKEGRSVVLTSHSMEC EALCTRMAIMVNGFRCLGSVOHLKNRFGD GYTITVRIAGSNPDLKPVQDFFGLAFPGSVPK EKHRNMLQYQLPSSLSSLARIFSILSQSKKRLH IEDYSVSQTTLDQVFVNFAKDQSDDDHLKDL SLHKNQTVVDVAVLTSFLQDEKVKESYV 10571 2327 A 8260 3 1567 PIPGTISFSLCFIPPCVPTMVRRPVSTISKGG YLQGNVNGRLPSLGNKEPPGQEKVQLKRKV TILRGVSIIIGTTIGAGIFISFKGVLQNTGSVOM SLTIMTVCGVLSLFGALSVAELGTTIKKSGGH YTYILEVFGPLPAFVRVWVELLIIRPAATAVIS LAFGRYILEPFFIQCEIPELAIKLITAVGITVVM VLNSMSVSWSARQIFLTFCKLTAILIIVPGV MQLIKGQTQNFKDAFSGROSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT FSERLLGNFSLAVIFIVALSCFGSMNGGVFAV SRLFYVASREGFILPELISMHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFSFARWLFIGLA VAGLIYLRVKCPDDHLPFFLAFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRILQILEVVPEEDKL PRINGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISESACEKQVHEVGLD GTETTYLPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FOPSLWPWDSVRNARSALTEMCVLTYDVLSIK KKSLAGAAQILLKGAERLTKSVTENQENKLQ ROFNSELLRLGHWKLKATEKLIGDLLSYRS AGSLPPHHGTEVIKNTDLDLDKKIPDSVCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSFHWQTKLEAAAQNVLLCKEI							AMADALERFGENRFVSPLSWDLVGRNLFAM
RRKRKPAVDRÍČVGIPPGEFGLLGVNGAGK SSTFKMLTGDTTVTRGDAFLNRNSILSNIHEV HQMMGYCPQFDAITELLTGREHVEFFALLRG VPEKEVGKVGEWAIRKLGLVKYGEKYAGNY SGGNKRKLSTAMALIGGPPVVFLDEPTTGMD PKARRFLWNCALSVVKEGRSVVLTSHSMEEC EALCTRMAIMVNGRFRCLGSVQHLKNRFGD GYTIVVRIAGSNPDILKPVQDFFGLAFPGSVPK EKHRNMLQYQLPSSLSSLARIFSILSQSKKRLH IEDYSVSGTILDQVFVNFAKDQSDDDHLKDL SLHKNQTVVDVAVLTSFLQDEKVKESYV 977 2327 A 8260 3 1567 IPGSTISFSLCFIPPPCVPTMVRKPVVSTISKGG YLQGNVNGRLPSLGNKEPPQQEKVQLKRKV TLLRGVSIIIGTIGAGIFISPKGVLQNTGSVGM SLTIWTVCGVLSLFGALSVAELGTTIKKSGGH YTYILEVFGPLPAFVRVWELLIRPAATAVIS LAFGRYILEPFFIQCEIPELAIKLITAVGITVVM VLNSMSVSWSARIQIFLITECKLTALILIIVPGV MQLIKGGTQNFKDAFSGRDSSITRLPLAFYGG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAELLLSNAVAVT SERLLGNFSLAVHFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVTILTGVPAYYLFII WDKKPRWFRIMSEKITRTLQILLEVVPEEDKL RINGCGLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVIESACEKGVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSL WPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILKKGABELIKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKLIGDLSYRS AGSLPPHHGTEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIGKQAPDIGDLGTVN LFKRPLPKSKPGSFHWQTKLEAAAQNVLLCKEI				!			AVEGVVFFLITVLIQYRFFIRPRPVNAKLSPLN
SSTFKMLTGDTTVTRGDAFINRNSILSNIHEV HQNMGYCPQFDAITELLTGREHVEFFALLRG VPEKEVGKVGEWARKLGLVKYGEKYAGNY SGGNKRLSTAMALIGGPPVFILDEPTTGMD PKARRFL WNCALSVVKEGRSVVLTSHSMEEC EALCTRMAIMVNGRFRCLGSVQHLKNRFGD GYTIVVRIAGSNPDLKPVQDFFGLAFPGSVPK EKHRNMLQYQLPSSLSSLARIFSILSQSKRLH IEDYSVSQTTLDQVFVNFAKDQSDDDHLKDL SLHKNQTVVDVAVLTSFLQDEKVKESYV 977 2327 A 8260 3 1567 IPGSTISFSLCFIEPFCVPTMVRRPVVSTISKGG YLQGNVNGRLPSLGNKEPPGGEKVQLKRKV TLLRGVSIIGTIBGGIFSPKGVLQNTGSVGM SLTIWTVCGVLSLFGALSYAELGTTIKKSGGH YTYILEVFGPLPAFVRVWVLLIIRPAATAVIS LAFGRYILEPFFQCEIPELAIKLITAVGITVVM VLNSMSVSWSARIQIFLTFCKLTAILIIVPGV MQLKGQTQNFKDAFSGROSSITRLPLAFYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLISNAVAVT PSERLLGNFSLAVIPVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRVKCPDMHRPFKVPLFFPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRILQILEVPEEDKL 978 2328 A 8261 2 2165 RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL GTETYLPPLSNSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FOPSL WPWDSVRNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKKARERLIKSVTENQENKLQ RDFNSELLRLRQHKIRKSVGEKIQDLSTVNL LFKRPLPKSKPGSFHQTIKLEAQNVLLCKEI DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSFHYQTKLEAAQNVLLCKEI							DEDEDVRRERQRILDGGGQNDILEIKELTKIY
HQNMGYCPQFDAITELLTGREHVEFFALLRG VPEKEVGKVGEWARKLGLVKYGEKYAGNY SGGNKRKLSTAMALIGGPPVVFLDEPTTGMD PKARRFLWNCALSVVEEGRSVVLTSHSMEEC EALCTRMAIMVNGFRFCLGSVQHLKNRFGD GYTIVVRIAGSNPDLKPVQDFFGLAFPGSVPK EKHRNMLQYQLPSSLSSLARIFSILSQSKKRLH IEDYSVSQTTLDQVFVNPAKDQSDDHLKDH SLHKNQTVVDVAVLTSFLQDEKVKESYV 977 2327 A 8260 3 1567 IPGSTISFSLCFIFPPCVPTMVRRPVVSTISKGG YLQGNVNGRLPGINEPPGQEKVQLKREV TLLRGVSIIIGTIIGAGIFISPKGVLQNTGSVGM SLTIWTVCGVLSLFGALSYAELGTTIKKSGGH YTVTLEVFGPLEAFALSVAELGTTIKKSGGH YTVTLEVFGPLEAFALSVAELGTTIKKSGGH WQLIKGQTQNFKDAFSGRDSSITRLPLAFTYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMHVRKHTPLPAIVL LHPLTMIMLFSGHVITLTGVPAYVLFII UPKKPRWFRIMSEKITRLQILEVYPEEDKL VAGLIYLRYKCPDMHRPFKVPLFIPALFSTC LFNVALSLYSDPFSTGIGFVITLTGVPAYVLFII WDKKPRWFRIMSEKITRLQILEVYPEEDKL PLADAASMSGVAVRISIESAGCEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWGAGSSADQDDEEGVVK FQPSLWFWDSVRNNLRSALTEMCVLYDVLSI VRDKKFRTLPRINGVASSADQDDEEGVVK KKSLAGAAQILLKGAGSSADQDDEEGVVK KKSLAGAAQILLKGAGSSADQDDEEGVVK KKSLAGAAQILLKGATLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI	1 1			1			RRKRKPAVDRICVGIPPGECFGLLGVNGAGK
977 2327 A 8260 3 1567 IPGSTISPLCHPPPCVFTMARPYOSTISKGG YLQGNVRGLSTAMALIGGPPVVFLDEPTTGMD PKARRFLWNCALSVVKEGRSVVLTSHSMECC EALCTRMAIMVNGFRECLGSVQHLKNRFGD GYTIVVRIAGSNPDLKPVQDFFGLAFPGSVPK EKHRNMLQYQLPSSLSSLARIFSILSQSKKRLH IEDYSVSQTTLDQVFVNPAKDQSDDDHLKDL SLHKNQTVVDVAVLTSFLQDEKVKESYV P77 2327 A 8260 3 1567 IPGSTISFSLCFIPPPCVPTMVRRPVVSTISKGG YLQGNVNGRLPSLGNEPPGQEKVQLKRKV TLLRGVSIIIGTIIGAGIFISPKGVLQNTGSVGM SLTIWTVCGVLSLFGALSYAELGTTIKKSGGH YTVILEVFGPLPAFVRVWELLIIRPAATAVIS LAFGRYILEPFFIQCEIPELAIKLITAVGITVM VLNSMSVSWSARQIFLTFCKLTAILIIIVPGV MQLIKGGTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVIEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMIHVRKHTPLPAVV LHFLTMMLFSGDLDSLLNFLSFARWLFFGLA VAGLIVLRYKCPDMHPPKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVTILTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL PNFWSLSSLSAGKPPLRTSFFGSWGVLPP LADAASMSGVAVRISIESACEKQWHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPFKQNPOTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLBLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI							SSTFKMLTGDTTVTRGDAFLNRNSILSNIHEV
SGGNKRLSTAMALIGGPPVVELDETTGMD PKARRFLWNCALSVVKEGRSVVLTSHSMEEC EALCTRMAMVNGRFRCLGSVQHLKNRFGD GYTIVVRIAGSNPDLKPVQDFFGLAFPGSVPK EKHRNMLQYQLPSSLSSLARIFSILSQSKKRLH IEDYSVSQTTLDQVFVNFAKDQSDDDHLKDL SLHKNQTVVDVAVLTSFLQDEKVKESYV 977 2327 A 8260 3 1567 IPGSTISFSLCFIFPPCVPTMVRKPVVSTISKGG YLQGNVNGRLPSLGNKEPPGQEKVQLKRKV TLLRGVSIIIGTIIGAGFISPKGVLQNTGSVGM SLTIWTVCGVLSLFGALSYAELGTTIKKSGGH YTYILEVFGPLPAFVRVWVELLIRPAATAVIS LAFGRYILEPFFIQCEIPELAIKLITAVGITVVM VLNSMSVSWARIQIFLTFCKLTAILIIIVPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNAYFTTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMYALSLYSDPFSTIGIGFVITLIGVPAYYLFII WDKKPRWFRIMSEKITRTLQHLEVVPEEDKL PKMYALSLYSDPFSTIGIGFVITLIGVPAYYLFII WDKKPRWFRIMSEKITRTLQHLEVVPEEDKL RHRGCGLLSSRLSAGKPPLRTSFFGSWOVLPY LADAASMSGVAVKISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVK FQPSLWFWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPOTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLIRLRQHWKLRKVGDKLGDLSYNS AGSLFPHIGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKFGSPHWQTKLEAAQNVLLCKEI]						HQNMGYCPQFDAITELLTGREHVEFFALLRG
PKARRFLWNCALSVVKEGRSVVLTSHSMEEC EALCTRMAIMVNGRFRCLGSVQHLKNRRGD GYTIVVRIAGSNPDLKPVQDFFGLAFPGSVPK EKHRNMLQYQLPSSLSSLARIFSILSQSKKRLH IEDYSVSQTTLDQVFVNFAKDQSDDHLKDI 977 2327 A 8260 3 1567 IPGSTISFSLCFIFPPCVPTMVRKPVVSTISKGG YLQGNVNGRLPSLGNKEPPGQEKVQLKRKV TLLRGVSIIIGTIGAGFISPKGVLQNTGSVGM SLTIWTVCGVLSLFGALSYAELGTTIKKSGGH YTYLLEVFGPLPAFVYWVELLIIRPAATAVIS LAFGRYILEPFTQCEIPELAIKLITAVGITVM VLNSMSVSWSARIQIFLTFCKLTAILIIIVPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPELLSMINVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHPPKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYVLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL PROMALSLYSDPFSTGIGFVITLTGVPAYVLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPGTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRIRQHWKLRKVGDKILGDLSYNR AGSLFPHIGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIGKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI	l i						VPEKEVGKVGEWAIRKLGLVKYGEKYAGNY
BALCTRMAIMVNGRFRCLGSVQHLKNRFGD GYTIVRIAGE GYTIVRIAGE EKHRYMLQYQLPSSLSSLARIFSILSQSKKRLH IEDYSVSQTTLDQVFVNFAKDQSDDDHLKDL SLHKNQTVVDVAVLTSFLQDEKVKESYV 977 2327 A 8260 3 1567 IPOSTISFSLEPPCVPTMYRKPVVSTISKGG YLQGNVNGRLPSLGNKEPPGQEKVQLKRKV TLLRGVSIIIGTIGAGIFISPKGVLQNTGSVGM SLTIWTVCGVLSLFGALSYABLGTTIKKSGGH YTYLLEVFGPLPAFVRVWVELLIRPAATAVIS LAFGRYILEPFFQCEIPBLAIKLITAVGITVVM VLNSMSVSWSARIQIFLTFCKLTAILIIVPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEBLLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKDPMHRPRKVPLFIPLAFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL PRAGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISESACEKQYHEVGLD GTETTYLPPLSMSQNLARLAQRIDFSQGSGSE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILKGAERLIKSVTENQENKLQ RDFNSELLRRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTEFVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKFGSPHWQTKLEAAQNVLLCKEI				1			SGGNKRKLSTAMALIGGPPVVFLDEPTTGMD
977 2327 A 8260 3 1567 IPGSTISFSLCFIFPPCVPTWARPVSTISKGG YLQGNVNORLPSLGNKEPPGEKVQLKRKV 977 2327 A 8260 3 1567 IPGSTISFSLCFIFPPCVPTWARPVSTISKGG YLQGNVNORLPSLGNKEPPGGEKVQLKRKV TLLRGVSIIIGTIIGAGIFISPKGVLQNTGSVGM SLTIWTVCGVLSLFGALSYAELGTTIKKSGGH YTYILEVFGPLPAFVRVWELLIIRPAATAVIS LAFGRYILEPFFIQCEIPELAIKLITAVGITVVM VLNSMSVSWSARGIFLTTCKLTAILIIIVPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMHVKKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL 978 2328 A 8261 2 2165 RGGSLRCVLGKLLGQILCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSGNLARLAQRIDFSQGSSEE EEAAGTEGDAQEWPGAGSSADQDDEGGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPFKQNPQTLQLISK KKSLAGAAQILLKGAERITKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIK VSIQKQAPDIGDLGTVN LFKRPLPRSKPGSPHWQTKLEAAQNVLLCKEI							PKARRFLWNCALSVVKEGRSVVLTSHSMEEC
SEKHRNMLQYQLPSSLSSLÄRIFSILSQSKKRLH IEDYSVSQTTLIDQVFVNFAKDQSDDDHLKDL SLHKNQTVVDVAVLTSFLQDEKVKESYV 977 2327 A 8260 3 1567 IPGSTISFSLCFIFPPCVPTMVRKPVVSTISKGG YLQGNVNGRLPSLGNKEPPGQEKVQLKRKV TLLRGVSIIGTIIGAGIFISPKGVLQNTGSVGM SLITIWTVCGVLSLFGALSYAELGTTIKKSGGH YTYILEVFGPLPAFVRVWVELLIRPAATAVIS LAFGRYILEPFFIQCEIPELAIKLITAVGITVVM VLNSMSVSWSARIQIFLTFCKLTAILIIIVPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLISNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSTTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRITQIILEVVPEEDKL VAGLIYLRYKCPDMHRPFKVPLFIPALFSTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRITQIILEVVPEEDKL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKPMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQIILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSANIKVSIQKQAPDIGDLGTVN LFKRPLPRSKPGSPHWQTKLEAAQNVLLCKEI							EALCTRMAIMVNGRFRCLGSVQHLKNRFGD
977 2327 A 8260 3 1567 IPGSTISFSLCFIPPPCVPTMVRKPVVSTISKGG YLQGNVNGRLPSLGNKEPPGQEKVQLKRKV TLLRGVSIIIGTIIGAGIFISPKGVLQNTGSVGM SLTIWTVCGVLSLFGALSYAELGTTIKKSGGH YTYILEVFGPLPAFVRVWVELLIIRPAATAVIS LAFGRYILEPFFIQCEIPELAIKLITAVGITVVM VLNSMSVSWSARIQIFLTFCKLTAILIIIVPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL 978 2328 A 8261 2 2165 RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTISFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPLSMSQNLARLAQRIDFSQGSGSE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI				,			GYTIVVRIAGSNPDLKPVQDFFGLAFPGSVPK
977 2327 A 8260 3 1567 IPGSTISFSLCFIPPCVPTMVRFVVSTISKGG YLQGNVNGRLPSLGNKEPPGQEKVQLKRKV TLLRGVSIIIGTIIGAGIFISPKGVLQNTGSVGM SLTIWTVCGVLSLFGALSYAELGTTIKKSGGH YTYILEVFGPLPAFVRVWVELLIIRPAATAVIS LAFGRYILEPFFQCEIPELAIKLITAVGITVVM VLNSMSVSWSARIQIFLTFCKLTAILIIIVPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV LHPLTMMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL PHANASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLVDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI							
977 2327 A 8260 3 IS67 IPGSTISFSLCFIFPPCVPTMVRKPVVSTISKGG YLQGNVNGRLPSLGNKEPPGQEKVQLKRKV TLLRGVSIIIGTIIGAGIFISPKGVLQNTGSVGM SLTIWTVCGVLSLFGALSYAELGTTIKKSGGH YTYILEVFGPLPAFVRVWVELLIIRPAATAVIS LAFGRYILEPFIQCEIPELAIKLITAVGITVVM VLNSMSVSWSARIQIFLTFCKLTAILIIIVPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPOTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI				[IEDYSVSQTTLDQVFVNFAKDQSDDDHLKDL
YLQGNVNGRLPSLGNKEPPGQEKVQLKRKV TLLRGVSIIIGTIIGAGIFISPKGVLQNTGSVGM SLTIWTVCGVLSLFGALSYAELGTTIKKSGGH YTVILEVFGPLPAFVRVWVEILLIIRPAATAVIS LAFGRYILEPFFIQCEIPELAIKLITAVGITVVM VLNSMSVSWSARIQIFLTFCKLTAILIIIVPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL PRIGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPOTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI							SLHKNQTVVDVAVLTSFLQDEKVKESYV
TLLRGVSIIIGTIIGAGIFISPKGVLQNTGSVGM SLTIWTVCGVLSLFGALSYAELGTTIKKSGGH YTYILEVFGPLPAFVRVWVELLIIRPAATAVIS LAFGRYILEPFFIQCEIPELAIKLITAVGITVVM VLNSMSVSWSARIQIFLTFCKLTAILIIIVPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTIEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLINFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL PRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTILQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELJRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI	977	2327	Α	8260	3	1567	IPGSTISFSLCFIFPPCVPTMVRKPVVSTISKGG
SLTIWTVCGVLSLFGALSYAELGTTIKKSGGH YTYILEVFGPLPAFVRVWVELLIRPAATAVIS LAFGRYILEPFFIQCEIPELAIKLITAVGITVM VLNSMSVSWSARIQIFLTFCKLTAILIIIVPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI							YLQGNVNGRLPSLGNKEPPGQEKVQLKRKV
YTYILEVFGPLPAFVRVWVELLIIRPAATAVIS LAFGRYILEPFFIQCEIPELAIKLITAVGITVVM VLNSMSVWSARIQIFLITFCKLTAILIIIVPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL 978 2328 A 8261 2 2165 RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI				}			TLLRGVSIIIGTIIGAGIFISPKGVLQNTGSVGM
LAFGRYILEPFFIQCEIPELAIKLITAVGITVVM VLNSMSVSWSARIQIFLTFCKLTAILIIIVPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL 978 2328 A 8261 2 2165 RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI	.						SLTIWTVCGVLSLFGALSYAELGTTIKKSGGH
VLNSMSVSWSARIQIFLTFCKLTAILIIIVPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRILQIILEVVPEEDKL 978 2328 A 8261 2 2165 RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI							YTYILEVFGPLPAFVRVWVELLIIRPAATAVIS
MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL 978 2328 A 8261 2 2165 RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI	1		İ				LAFGRYILEPFFIQCEIPELAIKLITAVGITVVM
MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRILQIILEVVPEEDKL 978 2328 A 8261 2 2165 RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI							VLNSMSVSWSARIQIFLTFCKLTAILIIIVPGV
MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRILQIILEVVPEEDKL 978 2328 A 8261 2 2165 RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI							MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG
FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL 978 2328 A 8261 2 2165 RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTILQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI							
FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL 978 2328 A 8261 2 2165 RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTILQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI							AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT
LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL 978 2328 A 8261 2 2165 RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI		j				}	
VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL 978 2328 A 8261 2 2165 RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI					!		SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV
VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL 978 2328 A 8261 2 2165 RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI							
978 2328 A 8261 2 2165 RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI							VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC
978 2328 A 8261 2 2165 RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI							LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII
978 2328 A 8261 2 2165 RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI							
RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI	978	2328	A	8261	2	2165	
LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI	İ	-					
GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI							- 1
EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI							
FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI					ł		, , ,
VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI					İ		
KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI							•
RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI							, , , ,
AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI					j		
DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI	[ſ	`
LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI						l	
THE PROPERTY OF THE PROPERTY O					}		FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ
LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE		-	ŀ			!	
HNLHLLIREFHKQTLSSIMMPHPASAPFGHKR		ļ	ŀ	. 1			
HILLILLIKEFHKQILSSIMMFHFASAFFGHKK							MAHULLIKEFUKQILSSIMIMIFUFASAFFUHKK

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion MRLSGPQAFDKNEINSLQSSEGLLEKIIKQAK HIFLRSRAAATIDSLASRIEDPQIQAHWSNIND VYESSVKVLITSQGYEQICKSIQLQLNIGVEQI RVVHRDGRVITLSYQEQELQDFLLSQMSQHQ VHAVQQLAKVMGWQVLSFSNHVGLGPIESIG NASAITVASPSGDYAISVRNGPESGSKIMVQF PRNQCKDLPKSDVLQDNKWSHLRGPFKEVQ WNKMEGRNFVYKMELLMSALSPCLL
979	2329	A	8289	2	1053	FVWNPRGGRKRRQAAVTQAATRASGTPSP RDGTMTQGKLSVANKAPGTEGQQQVHGEKK EAPAVPSAPPSYEEATSGEGMKAGAFPPAPTA VPLHPSWAYVDPSSSSSYDNGFPTGDHELFTT FSWDDQKVRRVFVRKVYTILLIQLLVTLAVV ALFTFCDPVKDYVQANPGWYWASYAVFFAT YLTLACCSGPRRHFPWNLILLTVFTLSMAYLT GMLSSYYNTTSVLLCLGITALVCLSVTVFSFQ TKFDFTSCQGVLFVLLMTLFFSGLILAILLPFQ YVPWLHAVYAALGAGVFTLFLALDTQLLMG NRRHSLSPEEYIFGALNIYLDIIYIFTFFLQLFG TNRE
980	2330	A	8305	59	857	ASQLPDYSISPPSLPPRISFHPSPTLARVAMAEP SEATQSHSISSSSFGAEPSAPGGGGSPGACPAL GTKSCSSSCAVHDLIFWRDVKKTGFVFGTTLI MLLSLAAFSVISVVSYLILALLSVTISFRIYKSV IQAVQKSEEGHPFKAYLDVDITLSSEAFHNY MNAAMVHINRALKLIIRLFLVEDLVDSLKLA VFMWLMTYVGAVFNGITLLILAELLIFSVPIV YEKYKTQIDHYVGIARDQTKSIVEKIQAKLPG IAKKKAE
981	2331	A	8308	186	1337	TRMSRHEGVSCDACLKGNFRGRRYKCLICYD YDLCASCYESGATTTRHTTDHPMQCILTRVD FDLYYGGEAFSVEQPQSFTCPYCGKMGYTET SLQEHVTSEHAETSTEVICPICAALPGGDPNH VTDDFAAHLTLEHRAPRDLDESSGVRHVRR MFHPGRGLGGPRARRSNMHFTSSSTGGLSSS QSSYSPSNREAMDPIAELLSQLSGVRRSAGGQ LNSSGPSASQLQQLQMQLQLERQHAQAARQ QLETARNATRTNTSSVTTITTQSTATTNIAN
982	2332	A	8315	1	1004	TESSQQTLQNSQFLLTRLNDPKMSETERQSM ESERADRSLFVQELLLSTLVREESSSSDEDDR GEMADFGAMGCVDIMPLDVALENLNLKESN KGNEPPPPPL GSTHASADAWAQWFCTEALVMGAPVWYLV AAALLVGFILFLTRSRGRAASAGQEPLHNEEL AGAGRVAQPGPLEPEEPRAGGRPRRRDLGS RLQAQRRAQRVAWAEADENEEAVILAQEE EGVEKPAETHLSGKIGAKKLRKLEEKQARKA QREAEEAEREERKRLESQREAEWKKEEERLR LEEEQKEEEERKAREEQAQREHEEYLKLKEA
983	2333	A	8320	244	1420	FVVEEEGVGETMTEEQSQSFLTEFINYIKQSK VVLLEDLASQVGLRTQDTINRIQDLLAEGTIT GVIDDRGKFIYITPEELAAVANFIRQRGRVSIA ELAQASNSLIAWGRESPAQAPA RRWRARGGLVPTLAWAEATGAYVPGRDKP DLPTWKRNFRSALNRKEGLRLAEDRSKDPHD PHKIYEFVNSGVGDFSQPDTSPDTNGGGSTSD TQEDILDELLGNMVLAPLPDPGPPSLAVAPEP CPQPLRSPSLDNPTPFPNLGPSENPLKRLLVPG EEWEFEVTAFYRGRQVFQQTISCPEGLRLVGS

	1 000 to	136	1.050	<u> </u>		
SEQ ID	SEQ ID	Met hod	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of nucl-	NO: of peptide	liou	ID NO:	beginning nucleotide	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
eotide	Seq-		USSN	location	location	F=Phenylalanine, G=Glycine, H=Histidine,
seq-	uence	[09/496	correspondi	corresponding to last amino	I=Isoleucine, K=Lysine, L=Leucine,
uence	dence		914	ng to first	acid residue	M=Methionine, N≈Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
dence			714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
]		peptide	Sequence	/=possible nucleotide deletion, \=possible
		1		sequence		nucleotide insertion
·		 	 	sequence		EVGDRTLPGWPVTLPDPGMSLTDRGVMSYV
						RHVLSCLGGGLALWRAGQWLWAQRLGHCH
	1			[1	TYWAVSEELLPNSGHGPDGEVPKDKEGGVF
		1		1		DLGPFIVGSLGPPDLITFTEGSGRSPRYALWFC
				1		VGESWPODOPWTKRLVMVKVVPTCLRALVE
		1				MARVGGASSLENTVDLHISNSHPLSLTSDQY
						KAYLQDLVEGMDFQGPGES
984	2334	A	8321	1	1243	ANMAPVEHVVADAGAFLRHAALQDIGKNIY
						TIREVVTEIRDKATRRRLAVLPYELRFKEPLPE
		1				YVRLVTEFSKKTGDYPSLSATDIQVLALTYOL
		İ	!			EAEFVGVSHLKQEPQKVKVSSSIQHPETPLHIS
						GFHLPYKPKPPQETEKGHSACEPENLEFSSFM
						FWRNPLPNIDHELQELLIDRGEDVPSEEEEEEE
		l			}	NGFEDRKDDSDDDGGGWITPSNIKQIQQELE
						QCDVPEDVRVGCLTTDFAMQNVLLQMGLHV
	İ		1			LAVNGMLIREARSYILRCHGCFKTTSDMSRV
			!			FCSHCGNKTLKKVSVTVSDDGTLHMHFSRNP
						KVLNPRGLRYSLPTPKGGKYAINPHLTEDQRF
		İ				PQLRLSQKARQKTNVFAPDYIAGVSPFVENDI
			ļ			SSRSATLQVRDSTLGAGRRRLNPNASRKKFV
						KKR
985	2335	Α	8322	352	529	RRNNIRQFIMKVCISGQARWLTPVVPVLWET
201						EAGRSLELKSLRPAWATWGNPISTKINK
986	2336	A	8325	89	1172	KMNPTDIADTTLDESIYSNYYLYESIPKPCTKE
						GIKAFGELFLPPLYSLVFVFGLLGNSVVVLVL
					·	FKYKRLRSMTDVYLLNLAISDLLFVFSLPFWG
						YYAADQWVFGLGLCKMISWMYLVGFYSGIF
						FVMLMSIDRYLAIVHAVFSLRARTLTYGVITS
						LATWSVAVFASLPGFLFSTCYTERNHTYCKT
			i i			KYSLNSTTWKVLSSLEINILGLVIPLGIMLFCY SMIIRTLQHCKNEKKNKAVKMIFAVVVLFLG
						FWTPYNIVLFLETLVELEVLQDCTFERYLDYA
						IQATETLAFVHCCLNPIIYFFLGEKFRKYILQL
						FKTCRGLFVLCQYCGLLQIYSADTPSSSYTQS
						TMDHDLHDAL
987	2337	A	8326	3	470	SLSAMRFLAATFLLLALSTAAQAEPVQFKDC
					.,	GSVDGVIKEVNVSPCPTQPCQLSKGQSYSVN
						VTFTSNIQSKSSKAVVHGILMGVPVPFPIPEPD
						GCKSGINCPIQKDKTYSYLNKLPVKSEYPSIK
						LVVEWQLQDDKNQSLFCWEIPVQIVSHL
988	2338	Α	8335	1205	323	VIKMALAARLLPQFLHSRSLPCGAVRLRTPA
	l	' I		ľ		VAEVRLPSATLCYFCRCRLGLGAALFPRSAR
1						ALAASALPAQGSRWPVLSSPGLPAAFASFPAC
					}	PQRSYSTEEKPQQHQKTKMIVLGFSNPINWV
			-			RTRIKAFLIWAYFDKEFSITEFSEGAKQAFAH
						VSKLLSQCKFDLLEELVAKEVLHALKEKVTS
ĺ	ĺ			ĺ		LPDNHKNALAANIDEIVFTSTGDISIYYDEKG
				!		RKFVNILMCFWYLTSANIPSETLRGASVFQVK
ŀ						LGNQNVETKQLLSASYEFQREFTQGVKPDWT
000	2005				10.0	IARIEHSKLLE .
989	2339	A	8349	67	185	MSGFIHQLLIQNLFCVYHTRLKTSQGLCLLSL
000	0340			210		KSLHPMS
990	2340	Α	8361	210	1115	ASPFLRPQGHDSGEREPFSQTPGLMQPFSIPVQ
]	1				ITLQGSRRRQGRTAFPASGKKRETDYSDGDPL
			Ì		ł	DVHKRLPSSTGEDRAVMLGFAMMGFSVLMF
	J	,		J		FLLGTTILKPFMLSIQREESTCTAIHTDIMDDW
					1	LDCAFTCGVHCHGQGKYPCLQVFVNLSHPG
				1		
			[1		QKALLHYNEEAVQINPKCFYTPKCHQDRNDL LNSALDIKEFFDHKNGTPFSCFYSPASQSEDVI

000	L OF STA		T 0= 0			
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq- uence	1	USSN 09/496	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq- uence	ucite		914	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
dence		1	914	ng to first amino acid	acid residue of peptide	Q=Glutamine, R=Arginine, S=Serine,
		1		residue of	sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
	Ì			peptide	sequence	/=possible nucleotide deletion, \=possible
ļ				sequence	·	nucleotide insertion
			 	Joquence		LIKKYDQMAIFHCLFWPSLTLLGGALIVGMV
	}	1	İ		· ·	RLTQHLSLLCEKYSTVVRDEVGGKVPYIEQH
	}		l	-		QFKLCIMRRSKGRAEKS
991	2341	A	8369	9	921	SSVVEFSALSVSMACLSPSQLQKFQQDGFLVL
			1207		'	EGFLSAEECVAMQQRIGEIVAEMDVPLHCRT
						EFSTQEEEQLRAQGSTDYFLSSGDKIRFFFEK
}	•	1	ļ		_	GVFDEKGNFLVPPEKSINKIGHALHAHDPVFK
i	İ					SITHSFKVQTLARSLGLQMPVVVQSMYIFKOP
ļ	ļ					HFGGEVSPHQDASFLYTEPLGRVLGVWIAVE
1			l			DATLENGCLWFIPGSHTSGVSRRMVRAPVGS
1		1				APGTSFLGSEPARDNSLFVPTPVQRGALVLIH
						GEVVHKSKQNLSDRSRQAYTFHLMEASGTT
						WSPENWLQPTAELPFPQLYT
992	2342	A	8370	906	4	MALSGNCSRYYPREQGSAVPNSFPEVVELNV
		' '	05,0	300	•	GGQVYFTRHSTLISIPHSLLWKMFSPKRDTAN
						DLAKDSKGRFFIDRDGFLFRYILDYLRDRQVV
	İ					LPDHFPEKGRLKREAEYFOLPDLVKLLTPDEI
1		Į				KQSPDEFCHSDFEDASQGSDTRICPPSSLLPAD
1						RKWGFITVGYRGSCTLGREGQADAKFRRVPR
1						ILVCGRISLAKEVFGETLNESRDPDRAPERYTS
		ĺ				RFYLKFKHLMGAPASNFILGFWGLGQNQDK
						HPVNIYLQQRSVIRPDLTSKKAGDLKGKGDA
		İ			. ,	QEVSRRRWLGDPEHL
993	2343	A	8379	1	2794	MRMQRHKNDTMDFGDSGKRIGGGVLCLLHQ
***	20 .5	l ''	03//	•	2//4	SNTSFIKLNNNGFEDIVIVIDPSVPEDEKIEQIE
		1	1			DMVTTASTYLFEATEKRFFFKNVSILIPENWK
1		1				ENPQYKRPKHENHKHADVIVAPPTLPGRDEP
						YTKQFTECGEKGEYIHFTPDLLLGKKQNEYG
			1			PPGKLFVHEWAHLRWGVFDEYNEDQPFYRA
						KSKKIEATRCSAGISGRNRVYKCQGGSCLSRA
						CRIDSTTKL YGKDCQFFPDKYQTEKASIMFM
						QSIDSVVEFCNEKTHNQEAPSLQNIKCNFRST
						WEVISNSEDFKNTIPMVTPPPPPVFSLLKIRORI
						VCLVLDKSGSMGGKDRLNRMNQAAKHFLLQ
						TVENGSWVGMVHFDSTATIVNKLIQIKSSDER
						NTLMAGLPTYPLGGTSICSGIKYAFQVIGELH
					J	SQLDGSEVLLLTDGEDNTASSCIDEVKQSGAI
						VHFIALGRAADEAVIEMSKITGGSHFYVSDEA
]				'	· • •	QNNGLIDAFGALTSGNTDLSQKSLQLESKGLT
			}	ŀ	1	LNSNAWMNDTVIIDSTVGKDTFFLITWNSLPP
						SISLWDPSGTIMENFTVDATSKMAYLSIPGTA
						KVGTWAYNLQAKANPETLTITVTSRAANSSV
						PPITVNAKMNKDVNSFPSPMIVYAEILQGYVP
					ļ	VLGANVTAFIESQNGHTEVLELLDNGAGADS
		ĺ				FKNDGVYSRYFTAYTENGRYSLKVRAHGGA
]						NTARLKLRPPLNRAAYIPGWVVNGEIEANPP
[[1		RPEIDEDTQTTLEDFSRTASGGAFVVSQVPSL
						PLPDQYPPSQITDLDATVHEDKIILTWTAPGD
				1		NFDVGKVQRYIIRISASILDLRDSFDDALQVN
						TTDLSPKEANSKESFAFKPENISEENATHIFIAI
1				ľ		KSIDKSNLTSKVSNIAQVTLFIPQANPDDIDPT
						PTPTPTPTPDKSHNSGVNISTLVLSVIGSVVIV
	,				ļ	NFILSTII
994	2344	A	8385	231	644	INSSPRTGRDHQELNLHTERDSRSQRAVLKIP
						RQNPGIFYWIFLPSRSHSASHGSRQRQVSCQG
]		TQDEILKMRNTFAELKNSLEALSSRMDQAEE
)]			j	1	J	RIGTQAGVQWRDHGSLQPQPPEFKQCFHLSL
Į	1				İ	PSSWDYRACLS
995	2345	Ā	8390	194	3421	AWRKSSVVPPRGTRRGEKSDQDKSGQKNKR
					J (4.1	TANIADOLATICATION OF POPULATION AND AND AND AND AND AND AND AND AND AN

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ĺ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		:	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
denoc			* * *	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	sequence	/=possible nucleotide deletion, \=possible
l			1	sequence		nucleotide insertion
		<u> </u>		scquence		
İ		ł				DFLSMKQSPALAPEERCRRAGSPKPVLRADD
						NNMGNGCSQKLATANLLRFLLLVLIPCICALV
	1	1	}			LLLEILLSYVGTLQKVYFKSNGSEPLVTDGEI
1						QGSDVILTNTIYNQSTVVSTAHPDQHVPAWT
						TDASLPGDQSHRNTSACMNITHSQCQMLPYH
ſ		[[ĺ	(ATLTPLLSVVRNMEMEKFLKFFTYLHRLSCY
						QHIMLFGCTLAFPECIIDGDDSHGLLPCRSFCE
	,					AAKEGCESVLGMVNYSWPDFLRCSQFRNQT
	į					ESSNVSRICFSPQQENGKQLLCGRGENFLCAS
I	1					GICIPGKLQCNGYNDCDDWSDEAHCNCSENL
ĺ			1	[FHCHTGKCLNYSLVCDGYDDCGDLSDEQNC
1						DCNPTTEHRCGDGRCIAMEWVCDGDHDCVD
]			KSDEVNCSCHSQGLVECRNGQCIPSTFQCDG
			ĺ			
						DEDCKDGSDEENCSVIQTSCQEGDQRCLYNP
				1		CLDSCGGSSLCDPNNSLNNCSQCEPITLELCM
1				1		NLPYNSTSYPNYFGHRTQKEASISWESSLFPA
ł				1		LVQTNCYKYLMFFSCTILVPKCDVNTGEHIPP
1						CRALCEHSKERCESVLGIVGLQWPEDTDCSQ
						FPEENSDNQTCLMPDEYVEECSPSHFKCRSGQ
ì						CVLASRRCDGQADCDDDSDEENCGCKERDL
1	1		į			WECPSNKQCLKHTVICDGFPDCPDYMDEKN
}				i		CSFCQDDELECANHACVSRDLWCDGEADCS
	!					DSSDEWDCVTLSINVNSSSFLMVHRAATEHH
	!					VCADGWQEILSQLACKQMGLGEPSVTKLIQE
						QEKEPRWLTLHSNWESLNGTTLHELLVNGQS
						CESRSKISLLCTKQDCGRRPAARMNKRILGGR
1						TSRPGRWPWQCSLQSEPSGHICGCVLIAKKW
ŀ	·					VLTVAHCFEGRENAAVWKVVLGINNLDHPS
1	1					VFMQTRFVKTIILHPRYSRAVVDYDISIVELSE
						DISETGYVRPVCLPNPEQWLEPDTYCYITGW
1						GHMGNKMPFKLQEGEVRIISLEHCQSYFDMK
1	ì					,
						TITTRMICAGYESGTVDSCMGDSGGPLVCEK
}						PGGRWTLFGLTSWGSVCFSKVLGPGVYSNVS
			00.00	100	0005	YFVEWIKRQIYIQTFLLN
996	2346	Α	8392	199	3085	KVILSSEMSKTNKSKSGSRSSRSRSASRSRSRS
						FSKSRSRSRSLSRSRKRRLSSRSRSRSYSPAHN
						RERNHPRVYQNRDFRGHNRGYRRPYYFRGR
		.				NRGFYPWGQYNRGGYGNYRSNWQNYRQAY
						SPRRGRSRSRSPKRRSPSPRSRSHSRNSDKSSS
						DRSRRSSSSRSSSNHSRVESSKRKSAKEKKSSS
						KDSRPSQAAGDNQGDEVKEQTFSGGTSQDTK
						ASESSKPWPDATYGTGSASRASAVSELSPRER
						SPALKSPLQSVVVRRRSPRPSPVPKPSPPLSST
						SQMGSTLPSGAGYQSGTHQGQFDHGSGSLSP
					İ	SKKSPVGKSPPSTGSTYGSSQKEESAASGGAA
						YTKRYLEEOKTENGKDKEOKOTNTDKEKIKE
						KGSFSDTGLGDGKMKSDSFAPKTDSEKPFRG
				}		SQSPKRYKLRDDFEKKMADFHKEEMDDQDK
						DKAKGRKESEFDDEPKFMSKVIGANKNQEEE
						KSGKWEGLVYAPPGKEKQRKTEELEEESFPE
						RSKKEDRGKRSEGGHRGFVPEKNFRVTAYK
1						AVQEKSSSPPPRKTSESRDKLGAKGDFPTGKS
						SFSITREAQVNVRMDSFDEDLARPSGLLAQER
						KLCRDLVHSNKKEQEFRSIFQHIQSAQSQRSP
					•	SELFAQHIVTIVHHVKEHHFGSSGMTLHERFT
	ĺ					KYLKRGTEQEAAKNKKSPEIHRRIDISPSTFRK
						HGLAHDEMKSPREPGYKAEGKYKDDPVDLR
			. 1			LDIERRKKHKERDLKRGKSRESVDSRDSSHSR
						ERSAEKTEKTHKGSKKQKKHRRARDRSRSSS
						SSSOSSHSYKAEEYTEETEEREESTTGFDKSRL
L			L			

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arghine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion GTKDFVGPSERGGGRARGTFQFRARGRGWG
						RGNYSGNNNNNSNNDFQKRNREEEWDPEYT PKSKKYYLHDDREGEGSDKWVSRGRGRGAF PRGRGRFMFRKSSTSPKWAHDKFSGEEGEIE DDESGTENREEKDNIQPTTE
997	2347	A	8398	202	552	CPALGGRQDLQGTRLLWAHDSGVGGQKAKS KQENLESLEATGREEEGGQGPPVTTKGVLLA LLMAGLALQPGTALLCYSCKAQVSNEDCLQ VENCTQLGEQCWTARIREWGDDSRQA
998	2348	A	8400	697	301	NPPSACTPGSCDSCSGRGRDLAFDSVWSTNN MSDPRRPNKVLRYKPPPSECNPALDDPTPDY MNLLGMIFSMCGLMLKLKWCAWVAVYCSFI SFANSRSSEDTKQMMSSFMLSISAVVMSYLQ NPOPMTPPW
999	2349	A	8401	93	1126	ASASHITSGHLRCFPGSEGVGTMARCFSLVLL LTSIWTTRLLVQGSLRAEELSIQVSCRIMGITL VSKKANQQLNFTEAKEACRLLGLSLAGKDQ VETALKASFETCSYGWVGDGFVVISRISPNPK CGKNGVGVLIWKVPVSRQFAAYCYNSSDTW TNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYS VASPYSTIPAPTTTPPAPASTSIPRRKKLICVTE VFMETSTMSTETEPFVENKAAFKNEAAGFGG VPTALLVLALLFFGAAAGLGFCYVKRYVKAF PFTNKNQQKEMIETKVVKEEKANDSNPNEES KKTDKNPEESKSPSKTTMRCLEAEV
1000	2350	A	8406	2	777	KERCQFVVKPMLSTVGSFLQDLQNEDKGIKT AAIFTADGNMISASTLMDILLMNDFKLVINKI AYDVQCPKREKPSNEHTAEMEHMKSLVHRL FTILHLEESQKKREHHLLEKIDHLKEQLQPLE QVKAGIEAHSEAKTSGLLWAGLALLSIQGGA LAWLTWWVYSWDIMEPVTYFITFANSMVFF
						AYFIVTRQDYTYSAVKSRQFLQFFHKKSKQQ HFDVQQYNKLKEDLAKAKESLKQARHSLCL QMQVEELNEKN
1001	2351	Á	8410	1400		VGFWERPLRSSRWFRRSLRRWEMLARAARG TGALLLRGSLLASGRAPRRASSGLPRNTVVLF VPQQEAWVVERMGRFHRILEPGLNILIPVLDR IRYVQSLKEIVINVPEQSAVTLDNVTLQIDGV LYLRIMDPYKASYGVEDPEYAVTQLAQTTM RSELGKLSLDKVFRERESLNASIVDAINQAAD CWGIRCLRYEIKDIHVPPRVKESMQMQVEAE RRKRATVLESEGTRESAINVAEGKKQAQILAS EAEKAEQINQAAGEASAVLAKAKAKAEAIRI LAAALTQHNGDAAASLTVAEQYVSAFSKLA KDSNTILLPSNPGDVTSMVAQAMGVYGALT KAPVPGTPDSLSSGSSRDVQGTDASLDEELDR VKMS
1002	2352	A	8421	134	941	NRENLLESRMMDPCSVGVQLRTTNECHKTY YTRHTGFKTLQELSSNDMLLLQLRTGMTLSG NNTICFHHVKIYIDRFEDLQKSCCDPFNIHKKL AKKNLHVIDLDDATFLSAKFGRQLVPGWKLC PKCTQIINGSVDVDTEDRQKRKPESDGRTAK ALRSLQFTNPGRQTEFAPETGKREKRRLTKN ATAGSDRQVIPAKSKVYDSQGLLIFSGMDLC DCLDEDCLGCFYACPACGSTKCGAECRCDRK WLYEQIEIEGGEIIHNKHAG
1003	2353	A	8427	3	1416	TEWGLSGSCPGCSPLEPGSRGRGAAAWRILR CRRLPEPSPFLTQPNLAQSQPPAPVPVTDPSVT MHPAVFLSLPDLRCSLLLLVTWVFTPVTTEIT

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first	Predicted end nucleotide location corresponding to last amino acid residue	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
				amino acid residue of peptide sequence	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						SLDTENIDEILNNADVALVNFYADWCRFSQM LHPIFEEASDVIKEEFPNENQVVFARVDCDQH SDIAQRYRISKYPTLKLFRNGMMMKREYRGQ RSVKALADYIRQQKSDPIQEIRDLAEITTLDRS KRNIIGYFEQKDSDNYRVFERVANILHDDCAF
						LSAFGDVSKPERYSGDNIIYKPPGHSAPDMVY LGAMTNFDVTYNWIQDKCVPLVREITFENGE ELTEEGLPFLILFHMKEDTESLEIFQNEVARQL ISEKGTINFLHADCDKFRHPLLHIQKTPADCP VIAIDSFRHMYVFGDFKDVLIPGKLKQFVFDL
1004	2354	A	8432	910	387	HSGKLHREFHHGPDPTDTAPGEQAQDVASSP PESSFQKLAPSEYRYTLLRDRDEL GLSRKLRAGFLPGFCRVSPCGSWVVETLVKM
1004	2334.	A	0432	910	36/	ACAAARSPADQDRFICIYPAYLNNKKTIAEGR RIPISKAVENPTATEIQDVCSAVGLNVFLEKN KMYSREWNRDVQYRGRVRVQLKQEDGSLC LVQFPSKKSVMLYAAEMIPKLKTRTQKTGGA
1005	2355	A	8453	90	530	DQSLQQGEGSKKGKGKKKK QSHETKMQSGTHWRVLGLCLLSVGVWGQD GNEEMGGITQTPYKVSISGTTVILTCPQYPGSE ILWQHNDKNIGGDEDDKNIGSDEDHLSLKEF SELEQSGYYVCYPRGSKPEDANFYLYLRARG NPGLQNRYHRLFREDHSKGHSQ
1006	2356	Α	8458	3	307	AVQRIRHEMNIFRLTGDLSHLAAIVILLLKIW KTRSCAGISGKSQLLFALVFTTRYLDLFTSFIS LYNTSMKVWYAIHRNVFHLQCTGLWTLNLC QLCIFN
1007	2357	Α	8459	43	553	GAGAGGDWAAMDKLKKVLSGQDTEDRSGL . SEVVEASSLSWSTRIKGFIACFAIGILCSLLGT VLLWVPRKGLHLFAVFYTFGNIASIGSTIFLM GPVKQLKRMFEPTRLIATIMVLLCFALTLCSA FWWHNKGLALIFCILQSLALTWYSLSFIPFAR DAVKKCFAVCLA
1008	2358	A	8462	487	150	AQDIRSVHSLGQKSTFVKHFRTLSHLHGLPDP PPHWPPQERSPPSHPCMPSHRPQIPQLSNSGPS DPRWGCVGPSMPTSTCLPGAVEASTTKASLP KCPVDSSLPTPEACFL
1009	2359	A	8465	134	954	ETRVKTSLELLRTQLEPTGTVGNTIMTSQPVP NETIIVLPSNVINFSQAEKPEPTNQGQDSLKKH LHAEIKVIGTIQILCGMMVLSLGIILASASFSPN FTQVTSTLLNSAYPFIGPFFFIISGSLSIATEKRL TKLLVHSSLVGSILSALSALVGFIILSVKQATL NPASLQCELDKNNIPTRSYVSYFYHDSLYTTD CYTAKASLAGTLSLMLICTLLEFCLAVLTAVL RWKQAYSDFPGSVLFLPHSYIGNSGMSSKMT HDCGYEELLTS
1010	. 2360	A	8468	2	473	KYRYRRPYPVMRKICQVGPAGLAFILNISPVA HRVALCHLAGCQEQAAWYHTLQILFFLVSAY FFSCPVPEKYFPGSCDIVGHGHQIFHAFLSICT LSQLEAILLDYQGRQEIFLQRHGPLSVHMACL SFFFLAACSAATAALLRHKVKARLTKKDS
1011	2361	A	8478	5	409	TELSQLEKAHPPADMGRRKSKRKPPPKKKMT GTLETQFTCPFCNHEKSCDVKMDRARNTGVI SCTVCLEEFQTPITCILGNLGFFQRVGRGLESG PCSSGPLCALVQGQSRPEEQVPPSDFCGVRRC RAGFQCQ
1012	2362	A	8481	2810	1652	RTSTQKWQSVFNDSQEHLERFYCNPENDRM RMKYGGQEFWADLNAMNVYETTEFDQLRR LSTPPSSNVNSIYHTVWKFFCRDHFGWREYPE

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion SVIRLIEEANSRGLKEVRFMMWNNHYILHNS FFRREIKRRPLFRSCFILLPYLQTLGGVPTQAP PPLEATSSSQIICPDGVTSANFYPETWVYMHP SQDFIQVPVSAEDKSYRIJYNLFHKTVPEFKYR ILQILRVQNQFLWEKYKKKEYMNRKMFGR DRIINERHLFHGTSQDVVDGICKHNFDPRVCG KHATMFQQGSYFAKKASYSHNFSKKSSKGV HFMFLAKVLTGRYTMGSHGMRRPPPVNPGS VTSDLYDSCVDNFFEPQIFVIFNDDQSYPYFVI
1013	2363	A	8488	2	517	QYEEVSNTVSI ENCRTRLRQAWHEVCGNKMAAPIPQOFSCL SRFLGWWFRQPVLVTQSAAIVPVRTKKRFTP PIYQPKFKTEKEFMQHARKAGLVIPPEKSDRS IHLACTAGIFDA YVPPEGDARISSLSKEGLIER TERMKKTMASQVSIRRIKDYDANFKIKDFPE KAKDIFIEGSPLY
1014	2364	Α .	8501	363	17	YIRTGYVYICIIYAQLMYTYYIRTAYVYICII.Y AQLMYTYVLYTHSLCIHMYSIRTAYVYICIIY AQIMYTYVFYTHRLCIHMYSIRTDYVYICILY AQLMYTYVFYTHSYMSDE
1015	2365	Α	8504	3	2190	NSSEHFSQAPQRLSFYSWYGSARLFRFRVPPD AVLLRWLLQVSRESGAACTDAEITVHFRSGA PPVINPLGTSFPDDTAVQPSFQVGVPLSTTPRS NASVNVSHPAPGDWFVAAHLPPSSQKIELKG LAPTCAYVFQPELLVTRVVEISIMEPDVPLPQ TLLSHPSYLKVFVPDYTRELLLELRDCVSNGS LGCPVRLTVGPVTLPSNFQKVLTCTGAPWPC RLLLPSPPWDRWLQVTAESLVGPLGTVAFSA VAALTACRPRSVTIQPLLQSSQNQSFNASSGL LSPSPDHQDLGRSGRVDRSPFCLTNYPVTRED MDVVSVHFQPLDRVSVRVCSDTPSVMRLRL
						NTGMDSGGSLTISLRANKTEMRNETVVVACV NAASPFLGFNTSLNCTTAFFQGYPLSLSAWSR RANLIIPYPETDNWYLSLQLMCPENAEDCEQ AVVHVETTLYLVPCLNDCGPYGQCLLLRHS YLYASCSCKAGWRGWSCTDNSTAQTVAQQR AATLLLTLSNLMFLAPIAVSVRFFLVEASVY AYTMFFSTFYHACDQPGEAVLCILSYDTLQY CDFLGSGAAIWVTILCMARLKTVLKYVLFLL GTLVIAMSLQLDRRGMWNMLGPCLFAFVIM ASMWAYRCGHRRQCYPTSWQRWAFYLLPG V\$MASVGIAIYTSMMTSDNYYYTHSIWHILL AGSAALLLPPPDQPAEPWACSQKFPCHYQIC KNDREELYAVT
1016	2366	A	8511	1	453	KWYPSGPVRIPGRFYYKLPAGHRRCRMAPAK KGGEKKKGRSAINEVVTREYTINIHKRIHGVG FKKRAPRALKEIRKFAMKEMGTPDVRIDTRL NKAVWAKGIRNVPYRIRVRLSRKRNEDEDSP NKLYTLVTYVPVTTFKNLQTVNVDEN
1017	2367	A	8513	54	1196	LERTPASADMAWTKYQLFLAGLMLVTGSINT LSAKWADNFMAEGCGGSKEHSFQHPFLQAV GMFLGEFSCLAAFYLLRCRAAGQSDSSVDPQ QPFNPLLFLPPALCDMTGTSLMYVALNMTSA SSFQMLRGAVIIFTGLFSVAFLGRRLVLSQWL GILATIAGLVVVGLADLLSKIIDSQHKLSEVIT GDLLIIMAQIIVAIQMVLEEKFVYKHNVHPLR AVGTEGLFGFVILSLLLVPMYYIPAGSFSGNP RGTLEDALDAFCQVGQQPLIAVALLGNISSIA FFNFAGISVTKELSATTRMVLDSLRTVVIWAL

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Trytophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion SLALGWEAFHALQILGFLILLIGTALYNGLHR PLLGRLSRGRPLAEESEQERLLGGTRTPINDA
1018	2368	A	8518	324	694	S SPFWTEKRRMEKPLFPLVPLHWFGFGYTALV VSGGIVGYVKTGSVPSLAAGLLFGSLAGLGA YQLYQDPRNVWGFLAATSVTFVGVMGMRS YYYGKFMPVGLIAGASLLMAAKVGVRMLM TSD
1019	2369	A	8526	2	1787	VSAAAVNMEPPDAPAQARGAPRLLLLAVLL AAHPDAQAEVRLSVPPLVEVMRGKSVILDCT PTGTHDHYMLEWFLTDRSGARPRLASAEMQ GSELQVTMHDTRGRSPPYQLDSQGRLVLAEA QVGDERDYVCVVRAGAAGTAEAAARLNVF AKPEATEVSPNKGTLSVMEDSAQEIATSNSRN GNPAPKITWYRNGQRLEVPVEMNPEGYMTS RTVREASGLLSLTSTLYLRKDDRDASFHC AAHYSLPEGRHGRLDSPTFHLTLHYPTEHVQ FWVGSPSTPAGWVREGDTVQLLCRGDGSPSP EYTLFRLQDEQEEVLNVNLEGNLTLEGVTRG QSGTYGCRVEDYDAADDVQLSKTLELRVAY LDPLELSEGKVLSLPLNSRAVVNCSVHGLPTP ALRWTKDSTPLGDGPMLSLSSITFDSNGTYVC EASLPTVPVLSRTQNFTLLVQGSPELKTAEIEP KADGSWREGDEVTLICSARGHPDPKLSWSQL GGSPAEPIPGRQGWVSSSLTLKVTSALSRDGI SCEASNPHGNKRHVFHFGTVSPQTSQAGVAV MAVAVSVGLLLLVVAVFYCVRRKGGPCCRQ RREKGAP
1020	2370	A	8530		1200	PRVRLLRPSRSRSCRGLLSTRAPGPSPFRSLHS SPLLPHAMKSPFYRCQNTTSVEKGNSAVMGG VLFSTGLLGNLLALGLLARSGLGWCSRRPLR PLPSVFYMLVCGLTVTDLLGKCLLSPVVLAA YAQNRSLRVLAPALDNSLCQAFAFFMSFFGL SSTLQLLAMALECWLSLGHPFFYRRHITLRLG ALVAPVVSAFSLAFCALPFMGFGKFVQYCPG TWCFIQMVHEEGSLSVLGYSVLYSSLMALLV LATVLCNLGAMRNLYAMHRRLQRHPRSCTR DCAEPRADGREASPQPLEELDHLLLALMTV LFTMCSLPVIYRAYYGAFKDVKEKNRTSEEA EDLRALRFLSVISIVDPWIFIIFRSPVFRIFFHKI FIRPLRYRSRCSNSTNMESSL
1021	2371	A	8536	1	237	RRGEIDMATEGDVELELETETSGPERPPEKPR KHDSGAADLERVTDYAEEKEIQSSNLETAMS VIGDRRSREQKAKQER
1022	2372	A	8537	94	541	RKERRRRRRMEAVVFVFSLLDCCALIFLSV YFIITLSDLECDYINARSCCSKLNKWVIPELIG HTIVTVLLLMSLHWFIFLLNLPVATWNIYRYI MVPSGNMGVFDPTEIHNRGQLKSHMKEAMI KLGFHLLCFFMYLYSMILALIND
1023	2373	A	8540	26	431	RMMKCPQALLAIFWLLLSWVSSEDKVVQSPL SLVVHEGDTVTLNCSYEVTNFRSLLWYKQEK KAPTFLFMLTSSGIEKKSGRLSSILDKKELSSIL NITATQTGDSAIYLCAVEAQCSLVTCSLYSNS TAEALQL
1024	2374	A	8544	1731	743	GVRLRYSPIAVVMVGEAGRDLRRRRAVAVT AEKMAVLAPLIALVYSVPRLSRWLAQPYYLL SALLSAAFLLVRKLPPLCHGLPTQREDGNPCD FDWREVEILMFLSAIVMMKNRRSITVEQHIGN IFMFSKVANTILFFRLDIRMGLLYITLCIVFLM

1025 2375 A 8546 2194 1707 TVSFHKTMASLKCSTVVCVICLEKPKYRCPA CRVPYCSVVCPRIKHKBQCNPETRPVEKIRS ALPITKTVKPVENNDDDDISADFLINSDEEDER VSLQNLKNLGESATLRSILLNPHLRQLMVKL DQGEDKAKLMRAYMQEPTSADCCLGIV VSLQNLKNLGESATLRSILLNPHLRQLMVKL DQGEDKAKLMRAYMQEPTSADCCLGIV EPSQNEES VGMELPAVVILKVILLGHWLLTTWGCIVFSGS YAWANFTILALGVWAVAQRDSDAISMFLGG LLATIFLDIVHISIFYRPVSLTDTGRFOVGMAII SLLKPLSCCFVYHMTRERGGELLVHTGFLG SSQDRSAYQTIDSABAPADPFAVPEGRSQDAF GY SSQDRSAYQTIDSABAPADPFAVPEGRSQDAF GY SSQDRSAYQTIDSABAPADPFAVPEGRSQDAF GY SSCDRSAYQTIDSABAPADPFAVPEGRSQDAF GY SFRYSGSEGSTOTLTKGELKYLMEKELPGFLG SKOKDAVDKLLKDLDANGDAQVDFSEFIV VAAITSACHKYFEKAGLK VAAATSACHKYFEKAGLK VAAATSACHKYFEKAGLK VAAATSACHKYFEKAGLK VAAATSACHKYFEKAGLK VAAATSACHKYFEKAGLK VAAATSACHKYFEKAGLK VAAT	SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
Seq. uence	1				1		F=Phenylalanine, G=Glycine, H=Histidine,
Unified				•			
maino acid tesidate of peptide residate of peptide residate of peptide sequence maintain per peptide sequence maintain per per per per per per per per per per						1	Q=Glutamine, R=Arginine, S=Serine,
Popsible nucleotide deterion, \=possible nucleotide deterion, \=possible nucleotide insertion				i	I		T=Threonine, V=Valine, W=Tryptophan,
	1					sequence	
1025 2375 A 8546 2194 1707 TVSPHENMENDE TVSPHENMEN			ļ	ļ			
WIVEFFANNSNDCQSFAPIYADLSLKYNCLET	ļ			 	sequence		
LNFGKVDVGRYTDVSTRYKVSTSPLTGCLFT LILEPGGKEAMRRPQDKKGRAVSWTSEED VIREPILNEL YQRAKKLSKAGDNIPEGPVAS TPTTVSDGENKKDK TPSTPLNEL YQRAKKLSKAGDNIPEGPVAS TPTTVSDGENKKDK TVSFHKTMASLKCSTVVCVICLEKPKYRCPA CRVPYCSVVCFRKHKBCQCNPETRPVEKKRS ALPIKTVKPVENDDDDSLADPLNSDEEDEDR VSLQNLKNLGESATLRSLLINPHRQDLWNLD DQGEDKAKLMRAYMQEPLFVEKKRS ALPIKTVKPVENDDDDSLADPLNSDEEDEDR VSLQNLKNLGESATLRSLLINPHRQDLWNLD DQGEDKAKLMRAYMQEPLFVERADCCLGIV EPSQNEES VGMELPAVVILKVILLGHWLLTTWCGIVPSGS YAWANFTILALGVWAVAQRDSIDAISMFLGG LLATIFLDIVHISIFYPRVSLTDTGRFCOYGMAII SLLLKPLSCCYYHMYRERGGELLVHTGFLG SSQDRSAYQTIDSABPADPFAVPEGRSQDAF GY GY GY STRYGSEGSTSTOTLTKGELLVHTGFLG SSQDRSAYQTIDSABPADPFAVPEGRSQDAF GY STRYGSEGSTSTOTLTKGELLVHMEKELPGFLK SKOKDAVDKLLKDLDANGDAQVDFSEFIV VAAITSACHKYFEKAGLK VAATSACHKYFEKAGLK LLLLGSCQGPQQVGAGQTFEYLKREHSLSKP YQGEAPRCFLRDWELQVFHKHGGKKNL HODGLAWYTKDRMQPGPVFGNMDKVGLC VFVDTYPNEEKQQEKVPYJSAMVNINGSLSY DHERDGRFTELGGCTAIVRNLHYDTFLVRY VKRHLTIMMDIDGKHE WRDCIEVPGVR.PRG YYFGTSSITGGLSIDHDVISLKLFELTVERTPE EKLHIDVPLFSVDNMKLPENTAPLPPLSGL ALFLIVFFSLVFSVFAIVIGIILTNKWQEQSKK RFY STRYSSEGSTELCLIPDLKGKARTREASSG AVASDPGSIL YPFIVVPITVTLLVTALVVWQS KRTEGIV SSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTISLCTSAKSSWTYRSGRLSWGIKC THLITTQALRQEHLSAGNGSHEPSSTYPSK VKRLEGIV VKRLEGUS VVSTVSTRYSGRLSWGIKC THLITTQALRQEHLSAGNGSCHEPSSTYPSK VKRLEGUS VKRLEGUS VKRLEGUS VFRIVVPITVTLLLVTALVVWQS KRTEGIV VKRLEGUS V			ĺ		ĺ	1	
LILFQGKEAMRRPQIDKKGRAVSWTFSERN VIRRFNLNEL YQRAKKLSKAGNPHEEQPVAS TPITIVSDGENKKDK TPITIVSDGENKKDK TRIFTVISDGENKKDK TVSFHKTMASLKCSTVCVICLEKPKYRCPA CRVPYCSVVCFRKHKEQCNPETRPVEKKIRS ALPIKTVKPVENKDDDDSIADFLINSDEEDR VSQNLKNLGSATLRSLLINFLRQLMVKLL DQGEDKAKLMRAYMQEPLFVEFADCCLGIV EPSQNEES POGRAVAN DQGEDKAKLMRAYMQEPLFVEFADCCLGIV EPSQNEES POGRAVAN POGRAV							LNFGKVDVGRYTDVSTRYKVSTSPLTKOLPT
1025 2375 A 8546 2194 1707				i	:		LILFQGGKEAMRRPQIDKKGRAVSWTFSEEN
1025	ļ				j	}	VIREFNLNELYQRAKKLSKAGDNIPEEQPVAS
CRVPYCSVVCFRKHKEQCNPETRPVEKKIRS	1005	0275	<u> </u>	9546	2104	1505	
ALPTKTYKPYENKDDDDSIADFINSDEEEDR VSLQNLKNI,GESATLASILLNPHLRQLMYNL DQGEDKAKLMRAYMQEPLFVEFADCCLGIV EPSQNEES YAWANFTILALGYWAVAQRDSIASMFLGG LLATIFLDIVHSIFYPRYSLITDTGRFGVGMAII SLLLKPLSCCFVYHMYRERGGELLVHTGFLG SQDRSAYQTIDSAEAPADPFAVPEGRQDAF GY GY DFLGPASPQEEGGSESTMTELETAMGMIDV VAAITSACHKYFEKAGLK VAATSACHKYFEKAGLK VAAITSACHK	1025	23/3	A	8546	2194	1707	CRYPYCENTICEPRITY CONDETERMENT IN
VSI,QNI,KNI,GESATLRSLLINPHLRQLMYNL	1		ļ				
DQGEDKARLMRAYMQEPLFVEFADCCLGIV				ļ	į]	
1026							
1027 2377 A 8557 1 340 DFLGPASPQEEGGSESSTMTELETAMGMIIDV 1028 2378 A 8569 20 963 KMAATLGPLGVWALUATIFLDIVALIVENCE/CYVENTYRERGGELLVITGFLG 1028 2378 A 8569 20 963 KMAATLGPLGSWQQWRGLSARDGSRRLLL 1028 LLLLGSQGPQQVGAGGTFEYLREHSLSRY 1028 VYAHTSACHKYPEKAGLK 1028 1028 1028 A 8569 20 963 KMAATLGPLGSWQQWRGLSARDGSRRLLL 1028 1028 1028 LLLLGSQGPQQVGAGGTFEYLREHSLSRY 1028 1029 1029 1029 1029 1029 1029 1029 1029					<u> </u>		
LLATIFLDIVHISIFYPRVSLIDTGRFGVGMAII SLLKKPLSCCFVYHMYRERGGELLVHTGFLG SSQDRSAYQTIIDSAAPADPFAVPEGRSQDAA GY DFLOPASPQEEGGSESSTMTELETAMGMIIDV FSRYSGSEGSTQTLTKGELKVLMEKELPGFLC SGKDKDAVDKLLKDLDANGDAQVDFSEFIVI VAAITSACHKYPEKAGLK 1028 2378 A 8569 20 963 KMAATLGFLGSWQQWRCLSARDGSRRLLL LLLLGSQGPQQVGAGQTFEYLKREHSLSKP YQGEAPRPCFLRDWELQVHFKHIGQGKKNL HGDGLAIWYTKDRMQPGPVFGNMDKFVGLG VFVDTYPNEEKQDEVTPYISAMVNNGSLSY DHERDGRPTELGGCTAIVRNLHYDTFLVIRY VKRHLTIMMDIDGKHEWRDCIEVPGVRLPRG YYFGTSSITGOLSDNHDVISLKLFELTVERTPE EEKLHRDVFLPSVDNMKLPEMTAPLPPLSGL ALFLIVFFSLVFSVFAIVIGILTYNKWQEQSRK RFY 1029 2379 A 8572 1 578 AAAASHRSRARSRPRRVSSGPAPRRAQSSAG RVASSGLDSAPLCTMARALCRLPRRGLWLLLA HHLPMTTACQEANYGALLRELCLTQFQVDM EAVGETL WCDWGRTIRSYRELADCTWHMAE KLGGCPWPNAEVDRFFLAVHGRYRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIU 1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPGGLCWSPRPLEK NKAMGRPLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEI AIVPNYRISWRRGHEHGQSFYSTRPPSHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTITT TTWRPSSTTTIAGGTVTESKGHSESWHISLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGITYPCFPKPGRTRHALCSVVLLLLTGQL	1026	2376	A	8547	1078	594	VGMELPAVNLKVILLGHWLLTTWGCIVFSGS
SLLLKPLSCCFVYHMYRERGGELLVHTGFLG SSQDRSAYQTIDSAEAPAPPAVPERRSQDAF GY	}		ł	1		1	YAWANFTILALGVWAVAQRDSIDAISMFLGG
1027 2377 A 8557 1 340 DFLGPASPQEEGGSESSTMTELETAMGMIIDV					1		
1027 2377 A 8557 1 340	1		ļ		 :		
1027 2377 A			}		i		
FSRYSGSEGSTQTLTKGELKVLMEKELPGFLC SGKDKDAVDKLKDLDANGDAQVDFSEFIVI VAAITSACHKYFEKAGLK	1027	2377	A	8557	1	340	
SGKDKDAVDKLLKDLDANGDAQVDFSEFIVI							FSRYSGSEGSTQTLTKGELKVLMEKELPGFLO
1028 2378 A 8569 20 963 KMAATLGPLGSWQQWRRCLSARDGSRRLLL LLLGSGGGPQQVGAGQTTEYLKREHSLSKP YQGEAPRPCFLRDWELQVHFKIHGQGKKNL HGDGLAIWYTKDRMQPGPVFGNMDKFVGLC VFVDTYPNEEKQQERVFPYISAMVNNGSLSY DHERDGRPTELGGCTAIVRNLHYDTFLVIRY VKRHLTIMMDIDGKHEWRDCIEVPGVRLPRG YYFGTSSITGDLSDNHDVISLKLFELTVERTPE EEKLHRDVFLPSVDNMKLPEMTAPLPPLSGL ALFLIVFFSLVFSVFAIVIGIILYNK WQEQSRK RFY 1029 2379 A 8572 1 578 AAAASHRSRARSRPRRVSSGPAPRRAQSSAG RVASGLDSAPLCTMARALCRLPRRGLWLLLA HHLFMTTACQEANY GALLRELCLTQFQVDM EAVGETLWCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV 1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKC THLTITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTITT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVALVLTVILGLLCLLLLWWRRKG SRAPSSDF RRTAGIYPCFFKFGRTRHALCSVVLLLLTGQL					ļ		SGKDKDAVDKLLKDLDANGDAQVDFSEFIVF
LLLLGSGQGPQQVGAGQTTEYLKREHSLSKP YQGEAPRPCFLRDWELQVHFKIHGQGKKNL HGDGLAIWYTKDRMQPGPVFGNMDKFVGLG VFVDTYPNEEKQQERVPPYISAMVNNGSLSY DHERDGRPTELGGCTAIVRNLHYDTFLVIRY VKRHLTIMMDIDGKHEWRDCIEVPGVRLPRG YYFGTSSITGDLSDNHDVISLKLFELTVERTPE EEKLHRDVFLPSVDNMKLPEMTAPLPPLSGL ALFLIVFFSLVFSVFAIVIGIILYNKWQEQSRK RFY 1029 2379 A 8572 1 578 AAASHRSRARSRPRRVSSGPAPRRAQSSAG RVASGLDSAPLCTMARALCRLPRRGLWLLLA HHLFMTTACQEANYGALLRELCLTQFQVDM EAVGETLWCDWGRTIRSYRELADCTWHMAE KLGGFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV 1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRISLCTSAKSSWTYRSGRLSSWGSIKG THLTITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGGSFYSTRPSHKDY VNRLFLNWTEGGESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVALVLKTVILGLLCLLLLWWRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFFKPGRTRHALCSVVLLLLTGQL					1		
VQGEAPRPCFLRDWELQVHFKIHGQGKKNL HGDGLAIWYTKDRMQPGPVFGNMDKFVGLC VFVDTYPNEEKQQERVFPYISAMVNNGSLSY DHERDGRPTELGGCTAIVRNLHYDTFLVIRY VKRHLTIMMDIDGKHEWRDCIEVPGVRLPRG YYFGTSSITGDLSDNHHDVISLKLFELTVERTPE EEKLHRDVFLPSVDNMKLPEMTAPLPPLSGL ALFLVFFSLVFSVFAIVIGIILYNKWQEQSRK RFY 1029 2379 A 8572 1 578 AAAASHRSRARSRPRRVSSGPAPRRAQSSAG RVASGLDSAPLCTMARALCRLPRRGLWLLLA HHLFMTTACQEANYGALLRELCLTQFQVDM EAVGETL WCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV 1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTISLCTSAKSSWTYRSGRLSWQSIKG THLITTQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEI AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFFKPGRTRHALCSVVLLLLTGQL	1028	2378	A	8569	20	963	, ,
HGDGLAIWYTKDRMQPGPVFGNMDKFVGLG VFVDTYPNEKQQERVFPYISAMVNNGSLSY DHERDGRPTELGGCTAIVRNLHYDTFLVIRY VKRHLTIMMDIDGKHEWRDCIEVPGVRLPRG YYFGTSSITGDLSDNHDVISLKLFELTVERTPE EEKLHRDVFLPSVDNMKLPEMTAPLPPLSGL ALFLIVFFSLVFSVFAIVIGIILYNKWQEQSRK RFY 1029 2379 A 8572 1 578 AAAASHRSRARSRPRVSSGPAPRRAQSSAG RVASGLDSAPLCTMARALCRLPRRGLWLLLA HHLFMTTACQEANYGALLRELCLTQFQVDM EAVGETLWCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV 1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFFKPGRTRHALCSVVLLLLTGQL							
VFVDTYPNEEKQQERVFPYISAMVNNGSLSY DHERDGRPTELGGCTAIVRNLHYDTELVIRY VKRHLTIMMDIDGKHEWRDCIEVPGVRLPRG YYFGTSSITGDLSDNHDVISLKLFELTVERTPE EEKLHRDVFLPSVDNMKLPEMTAPLPPLSGL ALFLIVFFSLVFSVFAIVIGIILYNKWQEQSRK RFY 1029 2379 A 8572 1 578 AAAASHRSRARSRPRVSSGPAPRRAQSSAG RVASGLDSAPLCTMARALCRLPRRGLWLLA HHLFMTTACQEANYGALLRELCLTQFQVDM EAVGETLWCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV 1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEI AIVPNVRISWRRGHFHQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFFKPGRTRHALCSVVLLLLTGQL							
DHERDGRPTELGGCTAIVRNLHYDTFLVIRY VKRHLTIMMDIDGKHEWRDCIEVPGVRLPRG YYFGTSSITGDLSDNHDVISLKLFELTVERTPE EEKLHRDVFLPSVDNMKLPEMTAPLPPLSGL ALFLIVFFSLVFSVFAIVIGIILYNKWQEQSRK RFY 1029 2379 A 8572 1 578 AAAASHRSRARSRPRRVSSGPAPRRAQSSAG RVASGLDSAPLCTMARALCRLPRRGLWLLLA HHLFMTTACQEANYGALLRELCLTQFQVDM EAVGETLWCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV 1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVKTVILGLLCLLLLWWRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL	1				ļ		
1029 2379 A 8572 1 578 AAAASHRSRARSRPRRVSSGPAPRRAQSSAG RVASGLDSAPLCTMARALCRLPRGLWLLLA HHLFMTTACQEANY GALLRELCLTQFQVDM EAVGETLWCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV 1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTLAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL						[
EEKLHRDVFLPSVDNMKLPEMTAPLPPLSGL							VKRHLTIMMDIDGKHEWRDCIEVPGVRLPRG
ALFLIVFFSLVFSVFAIVIGIILYNKWQEQSRK RFY 1029 2379 A 8572 1 578 AAAASHRSRARSRPRRVSSGPAPRRAQSSAG RVASGLDSAPLCTMARALCRLPRRGLWLLLA HHLFMTTACQEANYGALLRELCLTQFQVDM EAVGETLWCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV 1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL						1	I
1029 2379 A 8572 1 578 AAAASHRSRARSRPRRVSSGPAPRRAQSSAG RVASGLDSAPLCTMARALCRLPRRGLWLLLA HHLFMTTACQEANYGALLRELCLTQFQVDM EAVGETLWCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV 1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL							
1029 2379 A 8572 1 578 AAAASHRSRARSRPRVSSGPAPRRAQSSAG RVASGLDSAPLCTMARALCRLPRRGLWLLLA HHLFMTTACQEANYGALLRELCLTQFQVDM EAVGETLWCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV 1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEI AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL				}			
RVASGLDSAPLCTMARALCRLPRRGLWLLLA HHLFMTTACQEANYGALLRELCLTQFQVDM EAVGETLWCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV 1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL	1029	2379	A	8572	1	578	
EAVGETLWCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV 1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEI AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL	İ						RVASGLDSAPLCTMARALCRLPRRGLWLLLA
KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV 1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEI AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL	1						HHLFMTTACQEANYGALLRELCLTQFQVDM
AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV 1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL							
KRTEGIV 1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEI AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL							
1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEI AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF							
SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL	1030	2380	A	8574	1352	372	
THLTITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL							SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG
SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEI AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL		•					THLTITQALRQPLHRAPLLPGQLCWSPRPLEK
AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL							NKAMGRPLLLPLLLLQPPAFLQPGGSTGSGP
VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL	[
CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL		.					ALVEN V KLOWKROHFHOQSFYSTRPPSIHKDY
TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL						ļ	
AIRVALAVAVLKTVILGLLCLLLLWWRRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL]						TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT
1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL							AIRVALAVAVLKTVILGLLCLLLLWWRRRKG
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1						<u> </u>	SRAPSSDF
	1031	2381	Α	8580	905	340	
AFDDFQESCAMMWQKYAGSRRSMPLGARIL]]		AFDDFQESCAMMWQKYAGSRRSMPLGARIL
FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV							FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV
EQLQSHPEAQEALGPPLNIHYLKLIDRENFVD VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW							EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI
HLDEVFLELKDGQQIPVFKLSGENGDEVKKE							HLDEVFLELKDGOOIPVFKLSGENGDFVKKE
	1032	2382	Ā	8593	2558	961	RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR
WAAGALGVAGLLCAVLGAVMIVMVPSLIKQ	L						

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Trytophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion QVLKNVRIDPSSLSFNMWKEIPIPFYLSVYFFD VMNPSEILKGEKPQVRERGPYVYREFRHKSNI TFNNNDTVSFLEYRTFQFQPSKSHGSESDYIV
1033	2383	A	8595	595	767	MPNILVLGAAVMMENKPMTLKLIMTLAFTTL GERAFMNRTVGEIMWGYKDPLVNLINKYFP GMFPFKDKFGLFAELNNSDSGLFTGFTGVQNI SRIHLVDKWNGLSKVDFWHSDQCNMINGTS GQMWPPFMTPESSLEFYSPEACRSMKLMYKE SGVFEGIPTYRFVAPKTLFANGSIYPPNEGFCP CLESGIQNVSTCRFSAPLFLSHPHFLNADPVL AEAVTGLHPNQEAHSLFLDIHPVTGIPMNCSV KLQLSLYMKSVAGIGQTGKIEPVVLPLLWFA ESGAMEGETLHTFYTQLVLMPKVMHYAQYV LLALGCVLLLVPVICQIRSQEKCYLFWSSKK GSKDKEAIQAYSESLMTSAPKGSVLQEAKL AHLPDTLLLPPHSPTVPTPKSFOCSOKACFSRS
1055	2505	' '	0575	373	,0,	FCLLLSLVSSSLVSLSLCPPLTQA
1034	2384	Α	8597	640	164	VITSCIIPFAFGLGVRASERLAEIDMPYLLKYQ PMMQTIGQKYCMDPAVIAGVLSRKSPGDKIL VNMGDRTSMVQDPGSQAPTSWISESQVFQTT EVLITRITELQRRFPTWTPDQYLRGGLCAYSG GAGYVRSSQDLSCDFCNDVLARAKYLKRHG F
1035	2385	A	8603	936	204	AMASTLEYSPSPLRRLVGPAAGFSRAARADL SWDPMAFFTGLWGPFTCVSRVLSHHCFSTTG SLSAIQKMTRVRVVDNSALGNSPYHRAPRCI HVYKKNGVGKVGDQILLAIKGQKKKALIVG HCMPGPRMTPRFDSNNVVLIEDNGNPVGTRI KTPIPTSLRKREGEYSKVLAIAQNFV
1036	2386	A	8606	1	562	PTRAHSFDLCCSPCRRRLLGREEAGEEPTSPV TQYLQPRSPEECKMFACAKLACTPSLIRAGSR VAYRPISASVLSRPEASRTGEGSTVFNGAQNG VSQLIQREFQTSAISRDIDTAAKFIGAGAATVG VAGSGAGIGTVFGSLIIGYARNPSLKQQLFSY AILGFALSEAMGLFCLMVAFLILFAM
1037	2387	A	8615	2		SPGPSLPESAESLDGSQEDKPRGSCAEPTFTDT GMVAHINNSRLKAKGVGQHDNAQNFGNQSF EELRAACLRKGELFEDPLFPAEPSSLGFKDLG PNSKNVQNISWQRPKDIINNPLFIMDGISPTDI CQGILGDCWLLAAIGSLTTCPKLLYRVVPRG QSFKKNYAGIFHFQIWQFGQWVNVVVDDRL PTKNDKLVFVHSTERSEFWSALLEKAYAKLS GSYEALSGGSTMEGLEDFTGGVAQSFQLQRP PQNLLRLLRKAVERSLMGCSIEVTSDSELES MTDKMLVRGHAYSVTGLQDVHYRGKMETLI RVRNPWGRIEWNGAWSDSAREWEEVASDIQ MQLLHKTEDGEFWMSYQDFLNNFTLLEICNL TPDTLSGDYKSYWHTTFYEGSWRTGSSAGGC RNHPGTFWTNPQFKISLPEGDDPEDDAEGNV VVCTCLVALMQKNWRHARQQGAQLQTIGFV LYAVPKEFQNIQDVHLKKEFFTKYQDHGFSEI FTNSREVSSQLRLPPGEYIIPSTFEPHRDADFL LRVFTEKHSESWELDEVNYAEQLQEEKVSED DMDQDFLHLFKIVAGEGKEIGVYELQRLNR MAIKFKSFKTKGFGLDACRCMINLMDKDGSG KLGLLEFKILWKKLKKWMDIFRECDQDHSGT LNSYEMRLVIEKAGIKLNNKVMQVLVARYA DDDLIIDFDSFISCFLRLKTMFTFFLTMDPKNT GHICLSLEQVLGEGWEGICRIAPACPSTPPPPS

NO: of nucle could	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Couriego Couriego			hod	1			
Sequence 19449 2384 2388 23	1	,	1		1		
1038 2388 A 8621 3 1494 RSRMARAPEGVILLEGE CATE OF PERIOR OF PE	1	1 .	1				
amino acid residue of sequence peptide residue of sequence peptide sequenc		donce					
residue of peptide sequence peptide sequ			ĺ	***			T=Threonine, V=Valine, W=Tryntonhan
Peptide					I .		
					peptide	'	/=possible nucleotide deletion, \=possible
EAL	L		<u>L</u>		sequence		nucleotide insertion
1038							SDVPGPASCPRLFPPWDLLPVSTVAADDHVGI
HHLEPNYUPGGRPVREPEDITYTISLK VIT.NISLISLAKEETILTSWIGDIWQUPKLYSKODF GGIETLR VPSEL VWLPEIVLENNIDGQFGVAY DANWL VYEGGSVTWLPPALTYRSVCA VETYTYF PPDWQNCSLIFREGTYTAGEVETAVDNIDG KTINKIDIDTEA YTENGEWAIDFGFGVTRAFT GGATDGFGETDVTSLIRRERLFYVNINITYFC LISGLVLLAYFLPAQAGGQKCTVSINVLLAQV VFLFILAQKIPETSLSVPLIGRETLEYMVVATLI VMNCVIVLINSGRIPTITHAMSPELRIFULL LIRLLGSPPPEAPRAASSPRRASSVGLLRAE ELILKKPRSELVFEGGRHRQGTWTAAFCQS GAAAPEVRCCVJAVNIVVASERIDQEATGGE VSDWVRMGNALDNICFWAALVLFSVGSSLIF LGAYFNRYPDLPYAPCIQP LGAYFNRYPDLPYAPCIQP OAMENLPYPITSSSACATSSTSGASSSGCN NSSSGGGRRTGFQISVYSGPDRQTVOVIQQ ALHRQPSTAQVIQOMYAAQQALINTA QSWNSAAASGLAQQAVLIQNTSSPALTASQA QMYLRAQMIFTSTATVAVVQPELGTGSPAR QPTPAQVQNLTLRTQQTPAAASSPTTTQPVL SGNNSAQAPAGSSSINLASSPAAASGLAQQAVLICTSF HERRIGGSGCLSTGMAPALKGRPKKKPCP PSLALKFTPGGSQPLPTA ASQLAFGGKLTSTPSRFDFQCGGRGAVTCCSF HERRIGGSGCLSTGMAPALKGRPKKKPCP RRDSFSCVKDSNNSDGKAVKKCEARSA LTRFKRNNNNCKVSNEERFKVAUGEECCAMP QAFLVALYKYMKERKKPTERIPLYGFQNLW TMFQAAQKLGCYETTARGWKHYDELGG NPGSTSAATCTRRHYERLLPYERPIKGEEDKP LPPTKPRQENSSQENENKTKVSGTKRIKHEIP KSKKKENAPKPQDAAEVSSGCPELTISS KSIEPELPAADMKKKIEGVQFFSAKPLASRVD PFKDNETDQSNSEK VAEGAGGKETLISS KSIEPELPAADMKKKIEGVQFFSAKPLASRVD PFKDNETDQSNSEK VAEGAGGKETLISS KSIEPELPAADMKKKIEGVQFFSAKPLASRVD PFKDNETDQSNSEK VAEGAGGKETLISS KSIEPELPAADMKKKIEGVQFFSAKPLASRVD PFKDNETDQSNSEK VAEGAGGKETLISS KSIEPELPAADMKKKIEGVQFFSAKPLASRVD PFKDNETDQSNSEK VAEGAGGKETLISS KSIEPELPAADMKKRIEGVQFFSAKPLASRVD PFKDNETDQSNSEK VAEGAGGKETLISS KSIEPELPAADMKKRIEGVGFSAKPLASRVD PFKDNETDQSNSEK VAEGAGGERELTISS KSIEPELPAADMKKRIEGVGFSAKPLASRVD PFKDNETDQSNSEK VAEGAGGERELTISS KSIEPELPAADMKKRIEGVGFSAKPLASRVD PFKDNETDQSNSEK VAEGAGGERELTISS KSIEPELPAADMKKRIEGVGFSAKPLASRVD PFKDNETDQSNSEK VAEGAGGERELTISS KSIEPELPAADMKKRIEGVGFSAKPLASRVD PFKDNETDQSNSEK VAEGAGGERELTISS TTO TATATATATATATATATATATATATATATATATAT			L				!
ISLNEKEETLTTSVWIGIDWQDYRLNYSKDDG GGIETLA VPSEL VWIPEIVLENNIDGGFGVAJ DAWNLYYEGGSYTWLPPALYBSVCAVEVTYF PEDWQNCSLIFESGYTYALEPVETTAYDNDG KTINKIDIDTEAYTENGEWAIDFCPGVIRRHH GGATDGPGETDVITYSLIRRRPLFYVNNIUPCY LISGLVLLAYFILAQAGGGCCTYSINNILAQT VFILFILAQKIPETSLSYPLLGRELIFYMVVAIIL VMNCVIVLINSQGIPTTHAMSPBLRIVLLEL LEPALLGSPPPEAPRAASPPRAASSYGLLLARA ELILKKPRSELVFEGGRHRQGTWAAFCQSL GAAAPEVRCCVDAVNFVAESTROQEATGE VSDWVMGMALDINGFWAALVLFSVGSSLIF LGAYFNRYPDLPYAPCIQP GARPEGGGARRPOHLPALLPSERPDCATL QAMENLELPYPHTSSSACATSSTSGASSSSGCI SOSSONSAQAPAGSSSINLASSPAAQLLRA QSVNSAAASGAQQAVLLGNISSPATASQO QMYLRAQMLIFTHATATVTQPEAAAQLLRA QSVNSAAASGAQQAVLIGNISSPATASQO QMYLRAQMLIFTHATATVTQPECHTGSPAR PFTPAQVQNLTLRTQQTPAAAASGFTTQPVL PSLALKFTTGGSGCLSTGMAPILKGRPKKKPCPQ RRDSSFGVKDSNNNSDGRAVAKVECEARSA LIKPKNNINCKKVNSNEEKPKVAUGECRADE QAFLVALYKYMKERKTPIERIPYLGFKQNILW TMFQAQKLGGYGTTTARROWKTKHEIP KSKKEKENAPKPQDAAFVSSEQEKEGETLISQ KSIPEPLPAADMKKKIEGYQEFSARRLASRVO LPIKKRNNINCKKVNSNEEKPKVAUGECRADE QAFLVALYKYMKERKTPIERIPYLGFKQNILW TMFQAQKLGGYGTTTARROWKTKHEIP KSKKEKENAPKPQDAASKOPLTSSALUDSKGE LPIKKRNNINCKKVNSNEEKPKVAUGECRADE QAFLVALYKYMKERKTPIERIPYLGFKQNILW TMFQAQKLGGYGTTTARROWKTKHEIP KSKKEKENAPKPQDAAFVSSEQEKEGETLISQ KSIPEPLPAADMKKKIEGYQEFSARRLASRVO PFKRONINCKYNSNEEKPKVAUGECRADE QAFLVALYKYMKERKTPIERIPYLGFKQNILW TMFQAQKLGGYGTTTARROWKTKHEIP KSKKEKENAPKPQDAASKOPLTSSALUDSKGE LPIKKRNNINCKYNNEEKPKVAUGECRADE QAFLVALYKYMKERKTPIERIPYLGFKQNILW TMFQAQKLGGYGTTTSRROWKTKHEIP KSKKEKENAPKPQDAAFVSSEQEKEGETLISQ KSIPEPLPAADMKKIEGYQEFSARRLASRVO PFKRONINCKYNNINCHTITORIUNGT TMFGAQKLGGYGTTTSRROWKTKHEIP KSKKEKENAPKPQDAAFVSSEQEKEGETLISQ KSIPEPLPAADMKKIEGYQEFSARRLASRVO PFKRONINCKYNNINCHTITORIUNGT TMFGAQKLGGYGTTTSRROWKTKHEIP KSKMERMTNCPPWQTILTTAR TMFGTARTITORIUNGT TMFGTATTTORIUNGT TMFGTATTTORIUNGT TMFGTATTTORIUNGT TMFGTATTTORIUNGT TMFGTATTTORIUNGT TMFGTATTTORIUNGT TMFGTATTTORIUNGT TMFGTATTTORIUNGT TMFGTATTTORIUNGT TMFGTATTTORIUNGT TMFGTATTORIUNGT TMFGTATTORIUNGT	1038	2388	Α	8621	3	1494	
GGIETLR VYSEL VWLPENINDOGGEGVAY DANNL YYEGGSVTWLPPAITYS VANCAVEVTYF PPDWQNCSLIFRSQTYMAEVEFTF AVDNOG KTINKIDIDTEAYTEN GEWALDFCPGVIRRHH GGATDGPGETDVITYSLITRREPLETYWNIVIPCV LISGLVLLAYFLPAQAGGGCTVSINVLLAQI VVLFILIAQRIBETSIS VYPLIGRELIFYMVVATILI VMNCVIVLINSQRIPTIHAMSPRLRIFIVLEL LELLGSPPPERAPAASPPRRASSYGLLERAE ELILKKPRSEL VFEGQRHRQGTWTAAFCQSI GAAAPVRCCVDANTVAESTSTDGEATGEE VSDWVRMGNALDNICFWAALVLFSVGSSLIF LGAYPRVPDLPYAPCIQP AMBRELPVPHTISSS ACATSSTSGASSSGLIF LGAYPRRYDPLPYAPCIQP AMBRELPVPHTSSSACATSSTSGASSSGLIF ALQQHISASAQLOSLAAVQQASLVSNRQGST SGSNVSAQAPAQSSSINLAASPAAAQLLNRA QSVNSAAASGIAQAVLIGHTSTATATSQA QMYLRAQMLIFTFTATVATVQPELGTGSPA PPTPAQVQNLTLRTQQTPAAAASGPTTQPVL PSLALKPTPGGSQPLPTPA ASQLAFGGKLTSTPRFDFQCGCRGAVTCCSF HEHRHOSGRCLSTGMAPPLKGRPKKKPCPQ RRDSSSGVKSONNSDGKAVAKVKCEARSA LTKPKNHNCKKVSNEERPEVATIGEECRADE QAFLVALYKYMKERKTPIERIPYLGFKQINLS TIMFQAQKLLGGYGTTTARRYCHLLPYERIPKGECRADE QAFLVALYKYMKERKTPIERIPYLGFKQINLS SGSNSTAATCTRRHYERLLPYERIPKGECRADE QAFLVALYKYMKERKTPIERIPTLGFKQINLS KSKREKENAPKPQDAAEVSSGQEKGQETTLS LPYKPRQGENSSGPENKTKVSTKIKHEIP KSKKREKAPKPQDAAEVSSGQEKGQETTLS KSKREKENAPKPQDAAEVSSGQEKGQETTLS LPYKPRKQGENSSGPENKTKVSTKIKHEIP KSKKREKAPKPQDAAEVSSGQEKGQETTLS SGSRLFFLKKKIEGYQGFSALVDSKQGS KLCCTFESSESPCASSPRLPHHTGHRWQTR MRRRMTNCPPWQTLDTAPT GIKARTTORALDYCVQAGMKMMEQMLKEKK LPDLSGSESLEFLKVDYNYNNSHASTAYSPP NTSLAFYPGVGIKALTHICTANISTDWGFSES LFVLYNSFAEPKERVSTVNYNSHASTAYSPP NTSLAFYPGVGIKALTHICTANISTDWGFSES LFVLYNSFAEPKERVSTVNYNSHASTAYSPP NTSLAFYPGVGIKALTHICTANISTDWGFSES LFVLYNSFAEPKERVSTVNYNSHASTAYSPP NTSLAFYPGVGIKALTHICTANISTDWGFSES LFVLYNSFAEPKERVSTVNYNSHASTAYSPP NTSLAFYPGVGIKALTHICTANISTDWGFSES LFVLYNSFAEPKERVSTRUKHLERMUCPHASEW ALNANLSTLEVLTKIDNYTLLDYSLISSPEIPE NYLDLNLKGVFYPLENLTDPPSPSPPVPVLPER NYLDLNLKGVFYPLENLTDPPSPSPPVPVLPER SNSMLYIGIAEVFFKSASFAHFTAGVFNVTLS TEESSHPVONSQCGGGAVISCASTAYSPP NTSLAFPPONSQCGGAVISCASTAYSPP NTSLAFPPONGGCGGAVISCASTAYSPP NTSLAFPPONGGCGGAVISCASTAYSPP NTSLAFPPONGGCGGAVISCASTAYSPP NTSLAFPPONGGCGGAVISCASTAYSPP NTSLAFPPONGGCGGAVISCASTAYSPP NTSLAFPPONGGCGGAVISCASTAYSPP NTSLAFPPONGGCGGAVISCASTAYSPP NTSLAFPPONG							
DANNLYYEGGSVTWLPPAIYRSVCAVEUTY PPEPWONCSLIPRSOTYNABEPTAVDNOD							
PFDWQNCSLIFRSQTYNAEEVEFTFAVDNDG KTNKIDDTEAYTENGEWADGVOVERHIH GGATDGPGETDVIYSLIRRKPLFYVINIIVPCV LISGLVLLAYFLPAQAGGGKCTVSINVLLAQT VFLFLIAQKIFETSLSVPLLGRFLIFVMVVATL VMNCVIVLNVSQRTPTTHAMSPRLHVLLEL LPALLGSPPPEAPRAASPRASSVGLLRAAS ELLIKKPRSELVPEQGRIRQGTWTAAFCQSL GAAAPEVRCVDAVNFVAESTROQEATGGE VSDWYRMGNALDNICFWAALVLFSVGSSLIF LGAYFNRYPDLPYAPCQG 1039 2389 A 8636 1 900 PGERFGGGGARRPQHPALLPSERPCATL QAMENELPYPHTSSACATSTSCASSSGCN NSSSGGGRTGPQISVYSGIPPROTVQVIQQ ALHRQPSTAAQVLQQMYAAQQQHLMLQTA ALQQQHLSSAQLQSLAAVQQQASLVSNRQGST SGSNVSAQAPAQSSSINLAASPAAAQLLNRA QSVNSAAASGAQQAVLLGNFPTAVVAVQPELGTGSPAR PPTPAQVQNLTILRTQQTPAAAGQTHAMLGTA ALQQQHLSSAQQAVLLGNLAXGPTTQPVL PSLALKPTPGGSQPLPTPA 1040 2390 A 8645 98 1388 ASQLAFGGLISTDSRAFGTGPQSV PSLALKPTPGGSQPLPTPA 1044 2390 A 8645 98 1388 ASQLAFGGLISTTSRSDFGGCGRGAVTCCSF HERHRQSGRCLSTGMAPNLKGRPTRKKRPCPQ RRDSPSGVKDSNNSDGKAVAVKCEARSA LTKPKNNINCKKVSNEEKPKVAIGEECRADE QAFLVALYKYMKERKTPIERPYLGFKQNLW TMPQAAQKLGGVETTTARRQWKHIYDELGG NPGSTSAATCTRRHYERLILPYERFIKGEBDKP LPPIKPRQGENSSONNENGKKHKVGGTKRKIKHEIP KSKKEKENAPKPQDAAEVSSEQEKEQETLISQ KSKEKENAPKPQDAAEVSSEQEKEQETLISQ KSKEKENAPKPQDAAEVSSEQEKEQETLISQ KSKEKENAPKPQDAAEVSSEQEKEQETLISQ KSKEKENAPKPQDAAEVSSEQEKEQETLISQ KSKEKENAPKPQDAAEVSSEQEKEQETLISQ KSKEKENAPKPQDAAEVSSEQEKEQETLISQ KSKEKENAPKPQDAAEVSSEQEKEQETLISQ KSKEKENAPKPQDAAEVSSEQEKEQETLISQ KSKEKENAPKPQDAAEVSSEQEKEQETLISQ KSKEKENAPKPQDAAEVSSEQEKEQETLISQ KSKEKENAPKPQDAAEVSSEQEKEQETLISQ KSKEKENAPKPQDAAEVSSEQEKEQETLISQ KSKEKENAPKPQDAAEVSSEQEKEQETLISQ KSKEKENAPKPQDAAEVSSEQEKEQETLISQ KSKEKENAPKPQDAAEVSSEQEKEQETLISQ KSKEKENAPKPQDAAGKMEMAMLKKKE LPDLSGSSELEFLKVDYNYNFSNIKISAFSP NTSLAFVPQVCIKAT.TNHGTANISTDWGFESP LFVLVNSEAEPMEKPLIKNINFSNIKISAFSP NTSLAFVPQVCIKAT.TNHGTANISTDWGFESP LFVLVNSEAEPMEKPLIKNINFSNIKISAFSPP NTSLAFVPQVCIKAT.TNHGTANISTDWGFESP LFVLDLNIKGVPPLENITDPTSPSFIVLERGF NYLDLNIKGVPPLENITDPTSPSFIVLEGF NYLDLNIKGVPPLENITDPTSPSFIVLEGF NYLDLNIKGVPPLENITDPTSPSFIVLEGF NYLDLNIKGVPPLENITDPTVSGCFILSGLISQ NSMLYGIGAEPFKSAAVSQDNVPALQPKVKKLER LILITATATSTSCAAVSQDNVPALQPGKKKKLKR 1042 2392 A 8672 538 1700 ARRESTSCAAVSQDNVPALQPGK	1						
KTNIKIDIDTEAYTENGEMADICFOCYRIRHIN GGATOGGEGTUVYSLIRRKPLFYVININIVPCV LISGLVLLAYFI.PAQAAGGKCTVSINVLLAQT VFLFLIAQKIPETSLSVPLLGRFILIFWWVATLI VMNCVIVLIVVSQRTPTITHAMSPRLRHVILLEI LPRILGSPPPEAPRAASPPRRASSVGLLRAE ELILKKPRSELVFEQQRIRGGTWTAARCQSL GAAAPEVRCCVDAVNIVAESTRDQEATIGE VSDWYRMGANALDINCFWAALVLFSVGSSLIF LGAYFARVPDLPYAPCIQP PGRERFGGGGARREPQHIPALLPSERPDCATTL QAMENELPVPHTSSSACATSSTSGASSSSGCN NSSSGGGRFTGPQISVYSGIPDRQTVQVIQQ ALHRQPSTAAGVLQMYAAQQHLMILQTA ALQQQHLSSAQLQSLAAVQQASLVSNRQST SGSNVSAQAPAQSSSINLASAPAAQLLINKA QSVNSAAASGIAQQAVLIGNTSSPALTASQA QMYLRAGMILFTTATVATVQPELGTGSPAR PPTPAQVONLTLRTQQTPAAAASCPTPTQPVL PSLALKPTPGGSQPLPTPA 1040 2390 A 8645 98 1388 ASQLAFGGKLTSTPSRDFQGCGRGAVTCCSF HERRHQSGRCLSTGMAPNLKGRPRKKKPCPQ RRDSPSGWKDSNNSDGKAVAWKVCEARSA LTTEPKNNINCKKVSNEEKPRVAIGECECRADE QAPLVALYKYMKERKTPIELIYFERGEBOKP PERDIETTIVARGORGVITARRQWKHIYDELGG NPGSTSAATCTRRHYERLIPYFRPIGGEBOKP LPPIKPRKQENSSQENEMKTKVSGTKRKHEIP KSKKEKENAPRQDAAEVSEGGEGTLISQ KSIPEPLPAADMKKKIEGYQEFSAKPLASRVOD PEKDNETTIOQSNSEKTVAEGGEGKGFTIPLSQ KSKEKEKENAPRQDAAEVSEGGEGTLISQ KSIPEPLPAADMKKKIEGYQEFSAKPLASRVOD PEKDNETTIOQSNSEKVAEGAGRGGTPTPLPSA PLAFKDSAUVPGASKQPLTSPSAUVDSKQES KLCCFTESPESEPQEASFPRLPHHTGHRWQTR MRRMTNCPPVQUTLPTAP MRRMTNCPVPQUTLPTAP NTSLAFVPQVGIKALTNHGTANISTDWGFESP LFVLNNSFAEPMEKPLRNINEMLCPIKASSPIP NTSLAFVPQVGIKALTNHGTANISTDWGFESP LFVLNNSFAEPMEKPLRNINEMLCPIKASSPIP NTSLAFVPQVGIKALTNHGTANISTDWGFESP LFVLNNSFAEPMEKPLRNINEMLCPIKASSPIP NTSLAFVPQVGIKALTNHGTANISTDWGFESP LFVLNNSFAEPMEKPLRNINEMLCPIKASSPIP NTSLAFVPQVGIKALTNHGTANISTDWGFESP LFVLNNSFAEPMEKPLRNINEMLCPIKASSPIP NTSLAFVPQVGIKALTNHGTANISTDWGFESP LFVLNNSFAEPMEKPLRNINEMLCPIKASSPIP NTSLAFVPQVGIKALTNHGTANISTDWGFESP LFVLNNSFAEPMEKPLRNINEMLCPIKASSPIP NTSLAFVPROVGIKALTNHGTANISTDWGFESP LFVLNNSFAEPMEKPLRNINEMLCPIKASSPIP NTSLAFVPROVGIKALTNHGTANISTDWGFESP LFVLNNSFAEPMEKPLRNINEMLCPIKASSPIP NTSLAFVPROVGIKALTNHGTANISTDWGFESP LFVLNNSFAEPMEKPLRNINEMLCPIKASSPIP NTSLAFVPROVGIKALTNHGTANISTDWGFESP LFVLNNSFAEPMEKPLRNINEMLCPIKASSPIP NTSLAFVPROVGIKALTNHGTANISTDWGFESP LFVLNNSFAEPMEKPLRNINEMLCPIKASSPIP NTSLAFVPROVGIKALTHARGHTERS							
GGATDGGETDVIYSLIGREKPLFYVINIVPCY LISGLVILAYPLPAQAGGGKCYSINVILAQT VFIFILAQKIPETSLSVPILGRFIIEVMIVATII VMNCVIVIANVQRTPTTHAMSPRILRIVILEL LPRILGSPPPEAPRAASPPRRASSVGILLRAE ELILKKPRSELVPEQQRIRQGTWTAAFCQSL GAAAPEVRCVDAVNFVAESTROQEATGEE VSDWVRMGNALDNICFWAALVLFSVGSSLIF LGAYFNRVPDLPYAPCQG 1039 2389 A 8636 1 900 PGEERFGGGGARRRPQHEPALLPSERPDCATL QAMENELPVPHTSSACATSTSGASSSSGCN NSSGGGRFTGPQISVYSGIPDRGTVQVQQ ALHRQPSTAAQVLQQMYAAQOHLMLQTT ALQQQHLSSAQLQSLAAVQQASLVSNRQGST SGSNVSAAASGAQQAVLICATSSPALTSQA QSVNSAAASGAQQAVLGINRAAASGATTOQTAVTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA							· · · · · · · · · · · · · · · · · · ·
LISGL/VLLAYFLPAQAGGQKCTVSINVLLAQT VFLPFLIAGKPETIS.SVPLAFILFVMVVATI. VANCVIVLNVSQRTPTTHAMSPRLHYVLLEL LPRLLGSPPPPEAPRAASPPRASSPRLASVGULLRAE ELILKKPRSELVFEGQRHRQSTVTAAFCQSL GAAAPEVRCCVDAVNFVAESTEDQSATGEE VSDWVRMGMALDNICPWALIVLFSVGSSLIF LGAYFNRVPDLPYAPCIQP PGREPRGGGARRPOPULPALLPSERPDCATL QAMENELPVPHTSSSACATSSTSGASSSSGCN NSSSGGGGRRTCPQIBVYSGPDRQTVQVIQQ ALHRQPSTAAQYLQQMYAAQQQHLMLQTA ALQQHLSSAQLQSLAAVQQASLVSNRQGST SGSNVSQAPAQSSSINIAYAQAQ QMYLRAQMILITPTATVATVQPELGTGSPAR PPTPAQVQNLTLRTQQTPAAAASCPTPTQPVL PSLALPPTGGSQPLPTPL PSLALPTPGGSQPLPTPL PSLALPTPGGSQPLPTPL 1040 2390 A 8645 98 1388 ASQLAFGGKLTSTPSRDFQCGGRGAVTCCSF HEHRHQSGRCLSTGMAPNLKGRPKKKRCPQ RRDSFSGVKDSNNNSDGKAVAKVKCEARSA LTEPKNNHNCKKVSNEEREKVAUGECRADE QAFLVALLYKYMKERKTPIERPPLGFKQNLW TMFQAAQKLGGYETITARQWKHIPDELGG NPGSTSAATCTRRHVERLILPYERPIKGEEDKP LPPIKPRQENSSQENENKTKVSGTRIKHEIP KSKREKENAPKPQDAAAVSSEQEKEQETLISQ KSKEKENAPKPQDAAKSSEQEKEQETLISQ KSKEKENAPKPQDAAKSSEQEKEQETLISQ KSKEKENAPKPQDAAKSSEQEKEQETLISQ KSKEKENAPKPQDAAKSSEQEKEQETLISQ KSKEKENAPKPQDAAKSSEQEKEQETLISQ KSKEKENAPKPQDAAKSSEQEKEQETLISQ KSKEKENAPKPQDAAKSSEQEKEQETLISQ KSKEKENAPKPQDAAKSSEQEKEQETLISQ KSKEKENAPKPQDAAKSSEQEKEQETLISQ KSKEKENAPKPQDAAKSSEQEKEQETLISQ KSKEKENAPKPQDAAKSSEQEKEQETLISQ KSKEKENAPKPQDAAKSSEQEKEQETLISQ KSKEKENAPKPQDAAKSSEQEKEQETLISQ KSKEKENAPKPQDAAKSSEQEKEQETLISQ KSKEKENAPKPQDAAKSSEQEKEQETLISQ KSKEKENAPKPQDAAKSSEQEKEQETLISQ KSKEKENAPKPQDAAKSSEQEKEQETLISQ KSKEKENAPKPQDAAKSSEQEKEQETLISQ KSKEKENAPKPQDAAKSSEQUEKEGETLISQ KSKEKENAPKPQDAAKSSEQUEKEGETLISQ KSKEKENAPKPQDAAKSSEQUEKEGETLISQ KSKEKENAPKPQDAAGNISTISALTSTURGPEN PSLAFTVPQVGIKALTNDHGTANISTISAFSP NTSLAFVPQVGIKALTNDHGTANISTISAFSP NTSLAFVPQVGIKALTNDHGTANISTISAFSP NTSLAFVPQVGIKALTNDHGTANISTISAFSP NTSLAFVPQVGIKALTNDHGTANISTISAFSP NTSLAFVPQVGIKALTNDHGTANISTISAFSP NTSLAFVPQVGIKALTNDHGTANISTISAFSP NTSLAFVPQVGIKALTNDHGTANISTISAFSP NTSLAFVPQVGIKALTNDHGTANISTISAFSP NTSLAFVPQVGIKALTNDHGTANISTISAFSP NTSLAFVPQVGIKALTNDHGTANISTISAFSP NTSLAFVPQVGIKALTNDHGTANISTISAFSP NTSLAFVPQVGIKALTNDHGTANISTISAFSP NTSLAFVPQVGIKALTNDHGTANISTISAFSP NTSLAFVPQVGIKALTNDHGTANISTIS			l	i :			
VFLFLIAGKIPETSLSVPLLGRILIFWVVATLL VMNCVIVLNVSQRIPTIHAMSPRLRHVLLEL LPRLLGSPPPPEAPRASSPRRASSVGLLIRAE ELILKKPRSEL VFEGQRHRQGTVTAAFCQSL GAAAPEVRCCVDAVNFVAESTRDQEATGEE VSDWVRMGNALDNICFWAALVLFSVGSSLIF LGAYFNRVPDLPYAPCIQE LGAYFNRVPDLPYAPCIQE LGAYFNRVPDLPYAPCIQE LGAYFNRVPDLPYAPCIQE LGAYFNRVPDLPYAPCIQE LGAYFNRVPDLPYAPCIQE LGAYFNRVPDLPYAPCIQE LGAYFNRVPDLPYAPCIQE LGAYFNRVAGALSISTSGASSSSGCN NSSSGGSGRPTGPQISVYSGIPDRQTVQVIQQ ALHRQPSTAAQYLQMYAAQQASLYSNRQGST SGSNVSAQAPAQSSSINLAASPAAQLINKA QSVNSAAASGLAQQAVLICHTSGPAR PPTPAQVQMLTLRTQQTPAAASGPTTGPVPV PSLALKPTFGGSQPLPTPA ASQLAFGGKLTSFRRFQCGGRGAVTCCSF HEHRHQSGRCLSTGMAPNLKGRPRKKRCPQ RNDSPSGVKDSNNNSDGKAVAKVCCARSA LTEPKNNHNNCKVSNSELERFVRVAIGEECRADE QAFLVALYKYMKERKTPIERIPYLGFKQNLW TMFQAAQKLGGYETTARKKEPP KSKKEKENARARPQDAASVSEQEKEQETLISQ KSIPEPLPAADMKKKIEGYQETSARVINSHORLW PEKNNETTQGSSISEV LAREAGEKGPTPPLPSA PLAPEKDSALVPGASKQPLTSPSALVDSKQES KLCCFTESPESPEQASPRLPHITTGHRWQTR KRKEKENARARPQDAASVSEQEKEQETLISQ KSIPEPLPAADMKKKIEGYQETSARLASRVD PEKNNETTQGSNSEX VAEEAGEKGPTPPLPSA PLAPEKDSALVPGASKQPLTSPSALVDSKQES KLCCFTESPESPEQASPRLPHITTGHRWQTR MRRRMTNCPPWQTILPTAP 1041 2391 A 8646 113 1492 LLQEMCTKTIFVUGCFLLWNLYVSSQTTY GIGARITORALDYGVQAGMKMEGMLKSKK LPDLSGSESLEFILK VDYVNYNSPSIKISAFSPT NTSLAFVPGVGKALTNSTUMGFESP LFVLYNSFAEPMERPILKNLEMLCHIASEVK ALNANLSTLEVLTKIDNYLLDYSLISSPEITE NYLDLNLKGVFYPLENLTDPPSPYPPVPLER SNSMLYGIGATYFTKSASFAHFTAGVTNVTLS TEEISNHFVQNSQLGFWJSRTAETYNLOPFM VRIMATEPPIINLQPGNSTLDIPSINLATION VRIMATEPPIINLQPGNSTLDIPSINLATION VRIMATEPPIINLQPGNSTLDIPSINLATION VRIMATEPPIINLQPGNSTLDIPSINLATION VRIMATEPPIINLQPGNSTLDIPSINLATION VRIMATEPPIINLQPGNSTLDIPSINLATION VRIMATEPPIINLQPGNSTLDIPSINLATION VRIMATEPPIINLQPGNSTLDIPSINLATION VRIMATEPPIINLQPGNSTLDIPSINLATION VRIMATEPPIINLQPGNSTLDIPSINLATION VRIMATEPPIINLQPGNSTLDIPSINLATION VRIMATEPPIINLQPGNSTLDIPSINLATION VRI							
VMNCVIVLNVSQRTPTTHAMSPRLRHVLEL	1						VFLFLIAQKIPETSLSVPLLGRFLIFVMVVATLI
LPRILGSPPPEAPRASSPRRASSVGLILRAE ELILKEPRESE VFEGGRINGTWTAACOSL GAAAPEVRCCVDAVNFVAESTRDQEATGEE VSDWVRMGNALDNICFWAALVLFSVGSSLIF UGAVFRRVPDLPYAPCIQP 1039 2389 A 8636 1 900 PGRERPGGGGARRPQHLPALLPSERPDCATI. QAMENELPVPHTSSSACATSSTSGASSSSGCN NSSSGGSGRPTGFQISVYSGIPBRQTVQVIQQ ALHRQPSTAAQVIQQMYAAQQQHLMLOTA ALQQHLSSAQLQSLAAVQQASLVSNRQGST SGSNVSAQAPAQSSSINILAASPAAQLINRA QSVNSAQAASAGIAQQAVILLGNTSSPALTASQA QMYLRAQMLIFITTATVATVQPELGTGSPAR PPTPAQVQMLILTRTQTPAAAASGPTPTQPVL PSILALKPTGGSQPLPTPA 1040 2390 A 8645 98 1388 ASQLAFGGKLTSTPSRDFQGGGGAVTCSF HEHRHQSGRCLSTGMAPNILKGRPRKKKPCPQ RRDSFSGVKDSNNNSDGKAVAKVKCEARSA LTRFKNNHNNCKKVSNEERPRVAIGEECRADE QAFLVALYKYMKERKTPIERIPVLGFKQNLW TMFQQAQKLGGYETITARQWKHIYDELGG NPGSTSAATCTRRHVERLILPYERPIKGEEDKP LPPEPRR QENSSQENENKTKVSGTKRIKHEIP KSKKEKENAPRPQDAAAVSSEQEREQETILISQ KSIPEPLPAADMKKKIEGYQETSARLASRVD PEKDNETDGGSNSEV VAEEAGEKGPTPPLPSA PLAPEKDSALVPGASKOPLTSPSAJVDSKQES KICCFTESPSEPQCASPRLPHITGIRRWQTR MRRIMTNCPPWQTILPTAP 1041 2391 A 8646 113 1492 LLQRCTKTTPVLWGCFLLWNLYVSSQTTYP GIGARITGRALDYGVQAGKMMEQMLKEKK LPDLSGSESLEFLKVDYVNYNSNISIKAFSEP NTSLAFVPGVGKALTNIGTANISTDWGFESP LFVLYNSFAEPMERPILKNINEMLCPILASEVK ALDANLSTLEVLTKUDNYLDSVLISSPETTE NYLDLNLKGVFYPLENLTDPPFSPVPPVLPER SNSMLTVITUSMLSVSEPITE NYLDLNLKGVFYPLENLTDPPFSPVPPVLPER SNSMLTVITUSMLSVSRAFITTAGVFVVTLIS TEEISINHEVQNSQGLGGVUSRTAEIVTLIGGPLVCSLS LNKFRILALPESNRSNIEVLRFENLSSLHFGVL PLANAKLQGFFLPNBFLFLFVNSDIEVLEGF LLISTDLKYETSSKQQPSFHVWEGLNLTSRQW RGKSAP LAPRIATEESKAAVSQDNVPALQPGKKKKLR							VMNCVIVLNVSQRTPTTHAMSPRLRHVLLEL
GAAAPEVRCCVDAVNIVASTATEDCEATGE		[ĺ	[]			LPRLLGSPPPPEAPRAASPPRRASSVGLLLRAE
VSDWVRMGNALDNICFWAALVLFSVGSSLIF LGAYFNRYPDLPYAPGRANT CAYFNRYPDLPYAPGRANT CAYFNRYPDLPYAPGRANT CAYFNRYPDLPYAPGRANT CAMENGER CAYFNRYPDLPYAPGRANT CAMENGER CAYFNRYPDLPYAPGRANT CAMENGER CAYFNRAM CAYFNRA	1					,	
LGAYFNRYPDLPYAPCIQP	1	i	l				•
1039 2389 A 8636 I 900 PGRERPGGGARRRPQHIPALIPSERPDCATI QAMENELPYPHTSSSACATSSTSGASSSSGCN NSSSGGSRPTGPQISVYSGIPDRQTVQVIQQ ALHRQPSTAAQYLQQMYAAQQQHLMLQTA ALQQHLSSAQIQAAVLQASIVSRQGST SGSNVSAQAPAQSSSINLAASPAAAQLINRA QSVNSAAASGIAQQAVLGMTSSPALTASQA QMYLRAQMLIPTPTATVATVQPELGTGSPAR PPTPAQVQNLTLRTQQTPAAAASGPTPTQPVL PSLALKPTPGGSQPLPTPA ASQLAFGGKLTSTPSRDFQGCGRGAVTCCSF HEHRHQSGRCLSTGMAPNLKGRPRKKKPCPQ RRDSFSQVKDSNNDSGAVAKVGCARSA LTKPKNNHNCKKVSNEEKPKVAIGECRADE QAFLVALYKYMKERKTPIERPYLGFKQNLW TMFQAAQKLGGYFTATRARQWKHYDELGG NPGSTSAATCTRRTVERLILPYERFIKGEEDKP LPPIKPRKQENSSQENENKTKVSGTKRIKHEIP KSKKEKENAPKPQDAAEVSSEQEKEQETLISQ KSIPEPLPAADMKKKIEGYQFFSAKPLASRVD PEKDNETDQGSNSEKVAEAGEKGPTPILPSA PLAPEKDSALVPGASKQPLTSPSALVDSKQES KLCCTTESPSEPQEASFPRLPHTIGHRWQTR MRRRMTNCPPWQITLPTAP LQEMCTKTTPVLWGCFLLWNLYVSSSQTIYP GIKARITQRALDYGVQAGMKMIEQMLKEKK LPDLSGSESLEFILKVDVNYNYFSNIKISAFSPP NTSLAFVPGVGIKALTNHIGTANISTDWGFESP LFVLVNSFAEPMERPILKNLMEMCLQFILBSEVK ALNANLSTLEVLTKIDNYTLLDYSLISSPEITE NYLDLAKGVFYPLENLTDPFSPSPPFVLFER SNSMLYIGIAEYFFKSASFAHTTAGVFNVTLS TEEISNHFVQNSQGLGNVLSRIAEIYILSQPFM VRIMATEPPINLIQPGNFTLDIPASIMMLTQPK NSIVETIVSMDFVASTSVGLVILGQRLVCSLS LNRFRLALPESNRSNIEVLRFSNLSSILHFGVL PLANKALQQGFPHVFLFLVRSDILLFQVL SLL LNRFRLALPESNRSNIEVLRFSNLSSILHFGVL PLANKALQQGFPHVFLFLFVRSDILLFQVL SLL LNRFRLALPESNRSNIEVLRFSNLSSILHFGVL PLANKALQQGFPHVFLFLFVNSDILLFQVL SLL LNRFRLALPESNRSNIEVLRFSNLSSILHFGVL PLANKALQQGFPHVFLFLFVNSDILSTEGVL PLANKALQQGFPHVFLFLFVNSDILLFGVL PLANKALQQGFPHVFLFFLFVNSDILSTEGVL PLANKALQQGFPHVFLFFLFVNSDILLFGVL PLANKALQQGFPHVFLFFLFVNSDILSTEGVL PLANKALQQGFPHVFLFFLFVNSDILSTEGVL PLANKALQQGFPHVFLFFLFVNSDILSTEGVL PLANKALQQGFPHVFLFFLFVNSDILSTEGVL PLANKALQQGFPHVFLFFLFVNSDILFGVL PLANKALQQGFPHVFLFFLFVNSDILFGVL PLANKALQQGFPHVFLFFLFVNSDILFGVL PLANKALQQGFPHVFLFFLFVNSDILFTGVL PLANKALQQGFPHVFLFFFNLSSILHFGVL PLANKALQQGFPHVFLFFLFVNSDILFTGVL PLANKALQQGFPHVFLFFLFVNSDILFTGVL PLANKALQQGFPHVFLFFLFVNSDILFTGVL PLANKALQQGFPHVFLFFLFVNSDILFTGVL PLANKALQQGFPHVFLFFLFVNSDILFTGVL PLANKALQQGFPHVFLFFFNSDILSSILHFGVL PLANKALQQGFPHVFLFFLFVNSDILFTGVL PLANKALQQGFPHVFLFF							
QAMENELPYPHTSSSACATSSTSGASSSSCN NSSSGGSGRPTGPGISVYSGIPDRGTVQVIQQ ALHRQPSTAAQYLQQMYAAQQQHLMILQTA ALQQQHLSSAQLQSLAAVQQASLVSNRQGST SGSNVSAQAPQSSSINLAASPAAQLLNRA QSVNSAAASGIAQQAVLLGNTSSPALTASQA QMYLRAQMLIFTPTATVATVQPELGTGSPAR PPTPAQVQNLTLRTQQTPAAAASGPTPTQPVL PSLALKPTPGGSQPLPTPA 1040 2390 A 8645 98 1388 ASQLAFGGKLTSTPSRDFQGGGGAVTCCSF HEHRAGSGRCLTSTGMAPNLKGRPKKKPCPQ RRDSFSGVKDSNNNSDGKAVAKVKCEARSA LTKPKNNHNCKKVSNEEKPKVAIGEECRADE QAFLVALYKYMKERKTPIERIPYLGFKQINLW TMFQAQKLGGYETITARRQWKHYDELGG NPGSTSAATCTRRHYERLILPYERFIKGEEDKP LPPIKPRKQENSSQENENKTKVSGTKRIKHEIP KSKKEKENAPKPQDAAEVSSQEKEQETLISQ KSIPEPLPAADMKKEIGGYQFFSAKPLASRVD PEKDNETDQGSNSEKVAEEAGEKGPTPPLPSA PLAPEKDSALVPGASKQELTSPSALVDSKQES KLCCTTESPESEPQSEPGSFTLSPHTHTGHRWQTR MRRRMTNCPPWQTLPTAP 1041 2391 A 8646 113 1492 LLQEMCTKTIPVLWGCFLLWNLYVSSSQTIYP GIKARITQRALDYGVQAGMKMIEGMLKEKK LPDLSGSESLEFLKVDVYNYNFSNIKISAFSEP NTSLAFVPGVGIKALTNHGTANISTDWGFESP LFVLYNSSAEPMERPILKNLNEMLCPILASEVK ALNANLSTLEVLTKIDNYTLLDYSLISSPEITE NYLDLNLKGVFYPELPELTDPPFSPPVPLPER SNSMLYIGIAEYFKSASFAHFTAGVFNVTLS TEEISNHFVQNSQCLGNVLSIKAIETIYLLGSPFM VRIMATEPPININLQFGNFTLDIPASIMMLTQPK NSIVETIVSMDFVASTSVGLVILGQRLVCSLS LNRFRLALPESNRSNIEVLRFBNLSSILHFULGYFM VRIMATEPPININLQFGNFTLDIPASIMMLTQPFK NSIVETIVSMDFVASTSVGLVILGQRLVCSLS LNRFRLALPESNRSNIEVLRFBNLSSILHFULGVFPFM VRIMATEPPININLQFGNFTLDIPASIMMLTQPFM VRIMATEPPININLQFGNFTLDIPASIMMLTQPFM VRIMATEPPININLQFGNFTLDIPASIMMLTQPFM VRIMATEPPININLQFGNFTLDIPASIMMLTQPFM VRIMATEPPININLQFGNFTLDIPASIMMLTQPFM VRIMATEPPININLQFGNFTLDIPASIMMLTQPFM VRIMATEPPININLQFGNFTLDIPASIMMLTQPFM VRIMATEPPININLQFGNFTLDIPASIMMLTQPFM VRIMATEPPININLQFGNFTLDIPASIMMLTQPFM VRIMATEPPININLQFGNFTLDIPASIMMLTQPFM VRIMATEPPININLQFGNFTLDIPASIMMLTQPFM VRIMATEPPININLQFGNFTLDIPASIMMLTQPFM VRIMATEPPININLGFTNINLSSLIHFGGL PLANAKALQQGFFPLYPHHKFIFLYNSDLIHFGGL PLANAKALQQGFFPLYPHHKFIFLYNSDLIHFGGL PLANAKALQQGFFPLYPHHKFIFLYNSDLIFGGL PLANAKALQQGFFPLYBFILSSLIHTGGL PLANAKALQQGFFPLYBFILSSLIHTGGL PLANAKALQQGFFPLYBFILSSLIHTGGL PLANAKALQQGFFPLYBFILSSLIHTGGL PLANACAGAR ARRITATRESKAAVSQDNVPALQPGKKKKLR	1030	2290	<u> </u>	9626	1	000	
NSSSGGSRTGPOISVYSGIPDROTVQVIQQ ALHRQPSTAAQYLQQMYAAQQQHLMLQTA ALQQQHLSAQQLAAQQQSLVSNRQGST SGSNVSAQAPAQSSSINJAASPAAQLLNRA QSVNSAAASGAQAQASLINASPAATQLLNRA QSVNSAAASGAQQAVLIGNTSSPALTASQA QMYLRAQMLIFTPTATVATVQPELGTGSPAR PPTPAQVQNLTLRTQQ IPBAAASGPTPTQPVL PSLALKPTPGGSQPLPTPA 1040 2390 A 8645 98 1388	1039	2309	^	0030	'	900	- !
ALHRQPSTAAQYLQMYAAQQQHLMLQTA ALQQQHLSSAQLQSLAAVQQASLVSNRQGST SGSNVSAQAPAQSSSINLAASPAAAQLLNRA QSVNSAQAPAQSSSINLAASPAAAQLLNRA QSVNSAQAPAQSSSINLAASPAAAQLLNRA QSVNSAQAPAQSSSINLAASPAAAQLLNRA QSVNSAQAPAQSSSINLAASPAAAQLLNRA QSVNSAQAPAQSSSINLAASPAAAQLLNRA QSVNSAQAPAQSSSINLAASPAAAQLLNRA QSVNSAQAPAQSSSINLAASPAAQLLNRA QSVNSAQASGIQQAVLLGNTSSPARTASQA QMYLRAQMLIFTTATVATVQPELGTGSPAR PPTPAQVQNLTLRTQQTPAAASCGPTPTQPVL PSLALKPTPGSQPLPTPA 1040 2390 A 8645 98 1388 ASQLAFGGKLTSTSPSRDFQGGGGAVTCCSF HEHRHQSGRCLSTGMAPNLKGRPRKKKPCPQ RRDSFSGVKDSNNNSDGKAVAKVKCEARSA LTKPKNNHNCKKVSNBEKPKVAUGEECRADE QAFLVALYKYMKERKTPIERIPYLGFKQINLW TMFQAAQKLGGVETITARQWKHIYDELGG NPGSTSAATCTRRHYERLILPYERFIKGEEDKP LPPIKPRKQENSSQENENKTKVSGTKRIKHEIP KSKKEKENAPKPQDAAEVSSEQEKEQETLISQ KSIPEPLPAADMKKKIEGYQEFSAKPLASRVD PEKDNETIQGSNSEKVAEEAGEKGPTPPLPSA PLAPEKDSALVPGASKQPLTSPSALVDSKQES KLCCFTESPESEPQEASPRLPHHTGHRWQTR MRRRMTNCPPWQITLPTAP 1041 2391 A 8646 113 1492 LLQEMCTKTIPVLWGCFLLWNLYVSSSQTIYP GIKARITQRALDYGVQAGMKMEQMLKEKK LPDLSGSESLEFLKVDYVNYNFSNIKISAFSFP NTSLAFVPGYGIKALTNHGTANISTDWGFESP LFVLYNSFAEPMEKPILKNLNEMLCPILASEVK ALNANLSTLEVLTKIDNYTLLDVSLISSPEITE NYLDLNLKGVFYPLENLTDPPFSVPFVLPER SNSMLYIGIAEYFFKSASFAHFTAGVFNVTLS TEEISNHFVQNSQGLGNVLSRIAEIYILSQPFM VRIMATEPPINLQPGNTLDIPASIMMLTOPK NSIVETIVSMDPVASTSVGLVILGQRLVCSLS LNRFRLALPESNRSNIEVLRFENILSSILHFGVL PLANAKLQGGFPLPNPHKFLFVNSDIEVLEGF LLISTDLKYETSSKQQPSFHVWEGLNLISRQW RGKSAP LISTDLKYETSSKQQPSFHVWEGLNLISRQW RGKSAP ARRIARTRESKAAVSQDNVPALQPGKKKKLR		•	[
ALQQOHLSSAQLQSLAAVQQASLVSNRQGST SGSNVSAQAPAQSSSINLAASPAAAQLLNRA QSVNSAAAGIAQQAVLIGNTSSPALTASQA QMYLRAQMLIFTPTATVATVQPELGTGSPAR PPTPAQVQNLTLRTQQTPAAAASGPTPTQPVL PSLALKPTPGGSQPLPTTA 1040 2390 A 8645 98 1388 ASQLAFGGKLTSTPSRDFQGCGRGAVTCCSF HEHRHQSGRCLSTGMAPNLKGRPRKKKPCPQ RRDSFSGVKDSNNNSDGKAVAKVKCEARSA LTKPKNNHNCKKVSNEEKPKVAIGEECRADE QAFLVALYKYMKERKTPIERIPYLGFKQNLW TMFQAAQKLGGVETITARRQWKHIYDELGG NPGSTSAATCTRRHYERLILPYERFIKGEEDKP LPPIKPRKQENSSQENENKTKVSGTKIKHEIP KSKKEKENAPKPQDAAEVSSEQEKEQETLISQ KSIPPIPAADMKKKIEGYQEFSAKPLASRVD PEKDNETDQGSNSEKVAEEAGEKGPTPPLPSA PLAPEKDSALVPGASKQPLTSPSALVDSKQES KLCCFTESPESEPQEASPPRLPHHTGHRWQTR MRRRMTNCPPWQITLPTAP 1041 2391 A 8646 113 1492 LLQEMCTKTIPVLWGCFLLWNLYVSSSQTTYP GIKARITORALDYGVQAGMKMIEQMLKEKK LPDLSGSESLEFLKVDYVNYNFSNIKISAFSPP NTSLAFVPGVGIKALTNHGTANISTDWGFESP LFVLYNSFAEPMEREPILKNLNEMLCPHIASEVK ALNANLSTLEVLTKIDNYTLLDYSLISSPEITE NYLDLNLKGVFYPLENLTDPPFSPVFPVLPER SNSMLYIGIAEVFPLENLTDPPFSPVFPVLPER SNSMLYIGIAEVFPLENLTDPPFSPVFPVLPER SNSMLYIGIAEVFPLENLTDPFSPVFPVLPER SNSMLYIGIAEVFPLENLTDPFSPVFPVLPER SNSMLYIGIAEVFPLENLTDPFSPVFVVLPER SNSMLYIGIAEVFRLSSILHFGVL PLANAKLQGGPPLPNPHKFLFVNSDIEVLEGF LLISTDLKYETISSKQQPSFHVWEGLNLISRQW RGKSAP 1042 2392 A 8672 538 170 ARRIARTRESKAAVSQDNVPALQPGKKKKLR	ľ	İ	ĺ				
SGSNVSAQAPÁQSSSINLAÁSPAAQLLNRA QSVNSAAASGIAQQAVLGNTSSPALTASQA QMYLRAQMULITPTATVATVQPELGTGSPAR PPTPAQVQNLTLRTQQTPAAAASGPTPTQPVL PSLALKPTPGGSQPLPTPA 1040 2390 A 8645 98 1388 ASQLAFGGKLTSTPSRDFQCCGRGAVTCCSF HEHRHQSGRCLSTGMAPNLKGRPRKKKPCPQ RRDSFSGVKDSNNNSDGKAVAKVKCEARSA LTKPKNNHNCKKVSNEEKPKVAIGEECRADE QAFLVALYKYMKERKTPIERIPYLGFKQNNLW TMFQAAQKLGGYETITARRQWKHIYDELGG NPGSTSAATCTRRHYERLILPYERPIKGEEDKP LPPIKPRKQENSSQEINKKTKVSGTKRIKHEIP KSKKEKENAPKPQDAAEVSSEQEKEQETLISQ KSIPEPLPAADMKKKIEGYQEFSAKPLASRVD PEXDNETIDQGSNSEK VAEEAGEKGPTPPLPSA PLAPEKDSALVBASKQPLTSPSALVDSKQES KLCCFTESPESEPQEASFPRLPHHTGHRWQTR MRRRNTINCPPWGITLFTAP 1041 2391 A 8646 113 1492 LLCBMCTKTIPVLWGCFLLWNLYVSSSQTIYP GIKARITQRALDYGVQAGMKMIEQMLKEKK LPPLSGSESLEFLKVDYVNYNFSNIKISAFSFP NTSLAFVPGVGKALTNHGTANISTDWGFESP LFVLYNSFAEPMEKPILKNLNEMLCPIIASEVK ALNANLSTLEVLTKIDNYTLLDYSLISSPEITE NYLDLNLKGVFYPLENLTDPFSPVPFVLPER SNSMLYIGIAEYFFKSASFAHFTAGVFNVTLS TEEISMFVQNSQGLGNVLSRIAEIYILSQPFM VRIMATEPPINLQFGNFTLDIPASIMMLTQPK NSIVETIVSMDFVASTSVGLVILGGRLVCSLS LNRFRLALPESNRSNIEVLRFONLSSILHFGVL PLANAKLQGFPLPNPHKFLFVNSDIEVLEGF LLISTDLKYETSSKQQPSFHVWEGLNLISRQW RGKSAP 1042 2392 A 8672 538 170 ARRIARTRESKAAVSQDNVPALQPGKKKKLR							
QSVNSAAASGIAQQAVILGNTSSPALTASQA QMYLRAQMLIFTPTATVATVQPELGTGSPAR PPTPAQVQNLTLRTQQTPAAAASGPTPTQPVL PSLALKPTPGGSQPLPTPA 1040 2390 A 8645 98 1388 ASQLAFGGKLTSTPSRDFQCCGRGAVTCCSF HEHRHPGGSGCLSTGMAPNLKGRPRKKPCPQ RRDSFSGVKDSNNNSDGKAVAKVKCEARSA LTKPKNHNCKKVSNEEKPKVAIGEECRADE QAFLVAL VXYMKERKTPIERPYLGFKQNLW TMFQAAQKLGGYETITARRQWKHIYDELGG NPGSTSAATCTRRHYERLILPYERIKGEEDKP LPPIKPRKQEDRSSQENENKTKVSGTKRIKHEIP KSKKEKENAPKPQDAAEVSSEQEKEQETLISQ KSIPEPLPAADMKKKIEGYQEFSAKPLASRVD PEKDNETDQGSNSEKVAEEAGEKGPTPLPSA PLAPEKDSALVPGASKQPLTSPSALVDSKQES KLCCFTESPESEPQEASFRLPHHTGHRWQTR MRRRMITNCPPWQITLPTAP 1041 2391 A 8646 113 1492 LLQEMCTKTIPVLWGCFLLWNLYVSSSQTIYP GIKARITQRALDYGVQAGMKMEQMLKEKK LPDLSGSESLEFLKVDYVNYNPSNIKISAFSPP NTSLAFVPGVGIKALTNHGTANISTDWGFESP LFVLYNSFAEPMERPILKNLNEMLCPHASEVK ALNANLSTLEVLKIDNYTLDYSLISSPEITE NYLDLNLKGVFYPLENLTDPPFSPVPFVLPER SNSMLYIGIAEYFFKASSFAHFTAGVFNVTLS TEEISNHFVQNSQGLGNVLSRIAEIYILSQPFM VRIMATEPPINLQPGNFTLDIPASIMMLTQPK NSIVETIVSMDFVASTSVGLVULGGRLVCSLS LNRFRLALPESNRSNIEVLRFENLSSILHFGVL PLANAKLQQGPPLPNPHKFLFVNSDIEVLEGF LLISTDLKYETSSKQQPSFHVWEGLNLISRQW RGKSAP 1042 2392 A 8672 538 170 ARRIARTRESKAAVSQDNVPALQPGKKKKLR	1						
1040 2390 A 8645 98 1388 ASQLAFGGKLTSTPSRDFQGCGRGAVTCCSF HEHRHQSGRCLSTGMAPNLKGRPRKKKPCPQ RRDSFSGVKDSNNNSDGKAVAKVKCEARSA LTKPKNNHCKKVSNEEKPKVAIGEGECRADE QAFLVALYKYMKERKTPIERIPYLGFKQINLW TMFQAAQKLGGYETITARRQWKHIYDELGG NPGSTSAATCRRHYERLILPYERPIKGEEDKP LPPIKPRKQENSSQENENKTKVSGTKRIKHEIP KSKKEKENAPKPQDAAEVSSEQEKEQETLISQ KSIPEPLPAADMKKKIEGYQEFSAKPLASRVD PEKDNETDQGSNSEKVAEEAGEKGPTPPLPSA PLAPEKDSALVPGASKQPLTSPSALVDSKQES KLCCFTESPESEPQEASFPRLPHHTGHRWQTR MRRRMTNCPPWQITLPTAP A 8646 113 1492 LLQEMCTKTIPVLWGCFLLWNLYVSSSQTIYP GIKARITQRALDYQVQAGMKMIEQMLKEKK LPPLSGSELEFLKVDYVNYNFSNIKISAFSFP NTSLAFVPGVGIKALTNHGTANISTDWGFESP LFVLYNSFAEPMEKPILKNLNEMLCPHIASEVK ALNANLSTLEVLTKIDNYTLLDYSLISSPEITE NYLDLNLKGVFYPLENLTDPPFSPVPFVLPER SNSMLYIGIAEYFFKSASFAHFTAGVFNVTLS TEEISNHFVQNSQGLGNVLSRIAEIVILSQPFM VRIMATEPPIINLQPGNFTLDIPASIMMLTQPK NSIVETIVSMDIVVASTSVGLVILGQRLVCSLS LNRFRLALPESNRSNIEVLRFENILSSILHFGVL PLANAKLQQGFPLPNPHKFLFVNSDIEVLEGF LLISTDLKYETSSKQQPSFHVWEGLNLISRQW RGKSAP 1042 2392 A 8672 538 170 ARRIARTRESKAAVSQDNVPALQPGKKKKLR		1					
1040 2390 A 8645 98 1388 ASQLAFGGKLTSTPSRDFQGCGRGAVTCCSF HEHRHAPGGRCLSTGMAPNLKGRPKKKPCPQ RRDSFSGVKDSNNNSDGKAVAKVKCEARSA LTKPKNNHNCKKVSNEERPKVAIGEECRADE QAFLVALYKYMKERKTPIERIPYLGFKQINLW TMFQAQKLGGYETITARRQWKHIYDELGG NPGSTSAATCTRRHYERLILPYERFIKGEEDKP LPPIKPRKQENSSQENESKYKVSGTKRIKHEIP KSKKEKENAPKPQDAAEVSSEQEKEQETLISQ KSIPEPLPAADMKKKIEGYQEFSAKPLASRVD PEKDNETDQGSNSEKVAEEAGEKGPTPPLPSA PLAPEKDSALVPGASKOPLTSPSALVDSKQES KLCCFTESPESEPQEASFPRLPHHTGHRWQTR MRRRMTNCPPWQITLPTAP 1041 2391 A 8646 113 1492 LLQEMCTKTIPVLWGCFLLWNLYVSSSQTIYP GIKARITQRALDYGYQAGMKMIEQMIKEKK LPDLSGSESLEFLKVDYVNYNFSNIKISAFSFP NTSLAFVPGVGIKALTNHGTANISTDWGFESP LFVLYNSFAEPMERPLIKNLNEML,CPILASEVK ALNANLSTLEVLTKIDNYTLLDYSLISSPEITE NYLDLNLKGVFYYEENLTDPPFSPVPFVLPER SNSMLYIGIAEYFFKSASFAHFTAGVFNVTLS TEEISNHFVQNSQGLGNVLSRIAEIYILSQPFM VRIMATEPPIINLQPGNFTLDIPASIMMLTQPK NSIVETIVSMDFVASTSVGLIGGRUCVSLS LNFFRLALPESNRSNIEVLRFENILSSILHFGVL PLANAKLQQGFPLPNPHKFLFVNSDIEVLEGF LLISTDLKYETSSKQQPSFHVWEGINLISRQW RGKSAP							QMYLRAQMLIFTPTATVATVQPELGTGSPAR
1040 2390 A 8645 98 1388 ASQLAFGGKLTSTPSRDFQGCGRGAVTCCSF HEHRHQSGRCLSTGMAPNLKGRPRKKKPCPQ RRDSFSGVKDSNNNSDGKAVAKVKCEARSA LTKPKNNHNCKKVSNEERPKVAIGEECRADE QAFLVALYKYMKERKTPIERIPYLGFKQINLW TMFQAAQKLGGYETITARRQWKHIYDELGG NPGSTSAATCTRRHYERLILPYBRFIKGEEDKP LPPIKPRKQENSSQENENKTKVSGTKRIKHEIP KSKKEKENAPKPQDAAEVSSEQEKEQETLISQ KSIPPLPAADMKKKIEGYQEFSAKPLASRVD PEKNNETDQGSNSEKVAEEAGEKGPTPPLPSA PLAPEKDSALVPGASKQPLTSPSALVDSKQES KLCCFTESPESEPQEASPRLPHHTGHRWQTR MRRMTNCPPWQITLPTAP GIKARITQRALDYGVQAGMKMIEQMLKEKK LPDLSGSESLEFLKVDYVNYNFSNIKISATSFP NTSLAFVPGVGIKALTNHGTANISTDWGFESP LFVLYNSFAEPMEKPILKNLNEMLCPILASEVK ALNANLSTLEVLTKIDNYTLLDYSLISSPEITE NYLDLNLKGVFYPLENLTDPPFSPVPFVLPER SNSMLYIGIAEVFFKSASFAHFTAGVFNVTLS TEEISNHFVQNSQGLGNVLSRIAEIYILSQPFM VRIMATEPPINLQPGNFTLDIPASIMMLTQPK NSIVETIVSMDFVASTSVGLVILGQRLVCSLS LNRFRLALPESNRSNIEVLRFENILSSILHFGVL PLANAKLQQGFPLPNPHKFLFVNSDIEVLEGF LLISTDLK YETSSKQQPSFHVWEGLNLISRQW RGKSAP GKSAP ARRACKLER GRUNDER GRU	1						
HEHRHQSGRCLSTGMAPNLKGRPRKKKPCPQ RRDSFSGVKDSNNNSDGKAVAKVKCEARSA LTKPKNNHNCKKVSNEEKPKVAIGEECRADE QAPLVALYKYMKERKTPIERIPYLGFKQINLW TMFQAAQKLGGYETITARRQWKHIYDELGG NPGSTSAATCTRRHYERLILPYERFIKGEEDKP LPPIKPKQENSSQENENKTKVSGTKRIKHEIP KSKKEKENAPKPQDAAEVSSEQEKEQETLISQ KSIPPPLPAADMKKKIEGYQEFSAKPLASRVD PEKDNETDQGSNSEK VAEEAGEKGPTPPLPSA PLAPERDSALVPGASKQPLTSPSALVDSKQES KLCCFTESPESEPQEASFPRLPHHTGHRWQTR MRRRMTNCPPWQTTLPTAP 1041 2391 A 8646 113 1492 LLQEMCTKTIPVLWGCFLLWNLYVSSSQTIYP GIKARITQRALDYGVQAGMKMEQMLKEKK LPDLSGSESLEFLKVDYVNYNFSNIKISAFSFP NTSLAFVPGVGIKALTNHGTANISTDWGFESP LFVLYNSFAEPMEKPILKNLNEMLCPHIASEVK ALNANLSTLEVLTKIDNYTLLDYSLISSPEITE NYLDLNLKGYPYPLENLTDPPFSPVPPVLPER SNSMLYIGIAEYFFKSASFAHFTAGVFNVTLS TEEISNHFVQNSQGLGNVLSRIAEIYILSQPFM VRIMATEPPINLQPGNFTLDIPASIMMLTQPK NSTVETIVSMDFVASTSVGLVILGQRLVCSLS LNFRFLALPESNRSNIEVLRFENLSSILHFGVL PLANAKLQQGFPLPNPHKFLFVNSDIEVLEGF LLISTDLKYETSSKQQPSFHVWEGLNLISRQW RGKSAP 1042 2392 A 8672 538 170 ARRIARTRESKAAVSQDNVPALQPGKKKKLR	1040	0000	ļ	0.51.5			PSLALKPTPGGSQPLPTPA
RRDSFSGVKDSNNNSDGKAVAKVKCEARSA LTKPKNNFNCKKVSNEEKPKVAIGEECRADE QAFLVALYKYMKERKTPIERIPYLGFKQINLW TMFQAAQKLGGYETITARQWKHIYDELGG NPGSTSAATCTRRHYERLILPYERFIKGEEDKP LPPIKPRKQENSSQENENKTKVSGTTKRIKHEIP KSKKEKENAPKPQDAAEVSSEQEKEQETLISQ KSIPEPLPAADMKKKIEGYQEFSAKPLASRVD PEKDNETDQGSNSEKVAEEAGEKGPTPPLPSA PLAPEKDSALVPGASKQPLTSPSALVDSKQES KLCCFTESPESEPQEASPRLPHTGHRWQTR MRRRMTNCPPWQITLPTAP I041 2391 A 8646 113 1492 LLQEMCTKTIPVLWGCFLLWNLYVSSSQTIYP GIKARITQRALDYGVQAGMKMEQMLKEKK LPDLSGSESLEFLKVDYVNYNFSNIKISAFSFP NTSLAFVPGVGIKALTNHGTANISTDWGFESP LFVLYNSFAEPMEKPILKNLNEMLCPIIASEVK ALNANLSTLEVLTKIDNYTLLDYSLISSPEITE NYLDLNLKGVFYPLENLTDPPFSPVPPVLPER SNSMLYIGIAEYFFKSASFAHFTAGVFNVTLS TEEISNHFVQNSQGLGNVLSRIAEIYILSQPFM VRIMATEPPINLQPGNFTLDIPASIMMLTQPK NSTVETIVSMDFVASTSVGLVILGQRLVCSLS LNRFRLALPESNRSNIEVLRFENILSSILHFGVL PLANAKLQGFPLPNPHKFLFVNSDIEVLEGF LLISTDLKYETSKQQPSFHVWEGLNLISRQW RGKSAP 1042 2392 A 8672 538 170 ARRIARTRESKAAVSQDNVPALQPGKKKKLR	1040	2390	A	8645	98	1388	ASQLAFGGKLTSTPSRDFQGCGRGAVTCCSF
LTKPKNNHNCKKVSNEEKPK VAIGEECRADE QAFL VAL YK YMKERK TPIERIPYLGFKQINLW TMFQAAQKLGGYETITARRQWKHIYDELGG NPGSTSAATCTRRHYERLILPYERFIKGEEDKP LPPIKPRKQENSSQENENKTK VSGTKRIKHEIP KSKKEKENAPKPQDAAEVSSEQEKEQETLISQ KSIPEPLPAADMKKKIEGYQEFSAKPLASRVD PEKDNETDQGSNSEK VAEEAGEKGPTPPLPSA PLAPEKDSAL VPGASKQPLTSPSAL VDSKQES KLCCFTESPESEPQEA SFPRLPHHTGHRWQTR MRRRMTNCPPWQITLPTAP MRRRMTNCPPWQITLPTAP LQEMCTKTIPVLWGCFLLWNLYVSSSQTIYP GIKARITQRALDYGVQAGMKMIEQMLKEKK LPDLSGSESLEFLK VDYVNYNFSNIKISAFSFP NTSLAFVPGYGIKAI.TNHGTANISTDWGFESP LFVLYNSFAEPMEKPILKNLNEMLCPILASEVK ALNANLSTLEVLTKIDNYTLLDYSLISSPEITE NYLDLNLKGVFYPLENLTDPPFSPVPFVLPER SNSMLYIGIAEYFKSASFAHFTAGVFNVTLS TEEISNHFVQNSQGLGNVLSRIAEIYILSQPFM VRIMATEPPIINLQPGNFTLDIPASIMMLTQPK NSIVETIVSMDIPVASTSVGLVILGQRLVCSLS LNRFRLALPESNRSNIEVLRFENILSSILHFGVL PLANAKLQQGFPLPNPHKFLFVNSDIEVLEGF LLISTDLKYETSSKQQPSFHVWEGLNLISRQW RGKSAP 1042 2392 A 8672 538 170 ARRIARTRESKAAVSQDNVPALQPGKKKKLR	_}						
QAFLVALYKYMKERKTPIERIPYLGFKQINLW TMFQAAQKLGGYETITARRQWKHIYDELGG NPGSTSAATCTRRHYERLILPYERFIKGEEDKP LPPIKPRKQENSSQENEINKTKVSGTKRIKHEIP KSKKEKENAPKPQDAAEVSSEQEKEQETLISQ KSIPEPLPAADMKKKIEGYQEFSAKPLASRVD PEKDNETDQGSNSEK VAEEAGEKGPTPPLPSA PLAPEKDSALVPGASKQPLTSPSALVDSKQES KLCCFTESPESEPQEASFPRLPHHTGHRWQTR MRRRMTNCPPWQITLPTAP LQEMCTKTIPVLWGCFLLWNLYVSSSQTIYP GIKARITQRALDYGVQAGMKMIEQMIKEKK LPDLSGSESLEFLKVDYVNYNFSNIKISAFSFP NTSLAFVPGVGIKALTNHGTANISTDWGFESP LFVLYNSFAEPMEKPILKNLNEMLCPIIASEVK ALNANLSTLEVLTKIDNYTLLDYSLISSPEITE NYLDLNIKGVFYPLENLTDPPFSPVPFVLPER SNSMLYIGIAEYFFKSASFAHFTAGVFNVTLS TEEISNHEVQNSQGLGNVLSRIAEIYILSQPFM VRIMATEPPIINLQPGNFTLDIPASIMMLTQPK NSITVETIVSMDFVASTSVGLVILGQRLVCSLS LNRFRLALPESNRSNIEVLRFENILSSILHFGVL PLANAKLQQGFPLPNPHKFLFVNSDIEVLEGF LLISTDLKYETSSKQQPSFHVWEGLNLISRQW RGKSAP 1042 2392 A 8672 538 170 ARRIARTRESKAAVSQDNVPALQPGKKKKLR							
TMFQAAQKLGGYETITARRQWKHIYDĖLGG NPGSTSAATCTRRHYERLILPYERPIKGEEDKP LPPIKPRKQENSSQENENKTKVSGTKRIKHEIP KSKKEKENAPKPQDAAEVSSEQEKEQETLISQ KSIPEPLPAADMKKKIEGYQEFSAKPLASRVD PEKDNETDQGSNSEK VAEEAGEKGPTPPLPSA PLAPEKDSALVPGASKQPLTSPSALVDSKQES KLCCFTESPESEPQEASFPRLPHHTGHRWQTR MRRRMTNCPPWQITLPTAP 1041 2391 A 8646 113 1492 LLQEMCTKTIPVLWGCFLLWNLYVSSSQTIYP GIKARITQRALDYGVQAGMKMIEQMLKEKK LPPLSGSESLEFLKVDYVNYNFSNIKISAFSFP NTSLAFVPGVGIKALTNHGTANISTDWGFESP LFVLYNSFAEPMEKPILKNLNEMLCPIIASEVK ALNANLSTLEVLTKIDNYTLLDYSLISSPEITE NYLDLNLKGVFYPLENLTDPPFSPVPFVLPER SNSMLVIGIAEYFFKSASFAHITAGVFNVTLS TEEISNHFVQNSQGLGNVLSRIAEIYILSQPFM VRIMATEPPIINLQPGNFTLDIPASIMMLTOPK NSIVETIVSMDFVASTSVGLVILGQRLVCSLS LNRFRLALPESNRSNIEVLRFENILSSILHFGVL PLANAKLQQGFPLPNPHKFLFVNSDIEVLEGF LLISTDLKYETSSKQQPSFHVWEGLNLISRQW RGKSAP 1042 2392 A 8672 538 170 ARRIARTRESKAAVSQDNVPALQPGKKKKLR							
NPGSTSAATCTRRHYERLILPYERFIKGEEDKP LPPIKPRKQENSSQENENKTKVSGTKRIKHEIP KSKKEKENAPKPQDAAEVSSEQEKEQETLISQ KSIPEPLPAADMKKKIEGYQEFSAKPLASRVD PEKDNETDQGSNSEKVAEEAGEKGPTPPLPSA PLAPEKDSALVPGASKQPLTSPSALVDSKQES KLCCFTESPESEPQEASFPRLPHHTGHRWQTR MRRMTNCPPWQITLPTAP LQEMCTKTIPVLWGCFLLWNLYVSSSQTIYP GIKARITQRALDYGVQAGMKMIEQMLKEKK LPDLSGSESLEFLKVDYVNYNFSNIKISAFSFP NTSLAFVPGVGIKALTNHGTANISTDWGFESP LFVLYNSFAEPMEKPILKNLNEMLCPHASEVK ALNANLSTLEVLTKIDNYTLLDYSLISSPEITE NYLDLNLKGVFYPLENLTDPPFSPVPFVLPER SNSMLYIGIAEYFFKSASFAHFTAGVFNVTLS TEEISNHFVQNSQGLGNVLSRIAEIYHLSQPFM VRIMATEPPHNLQPGNFTLDIPASIMMLTQPK NSIVETIVSMDFVASTSVGLVHGQRLVCSLS LNRFRLALPESNRSNIEVLRFENILSSILHFGVL PLANAKLQQGFPLPNPHKFLFVNSDIEVLEGF LLISTDLKYETSSKQQPSFHVWEGLNLISRQW RGKSAP 1042 2392 A 8672 538 170 ARRIARTRESKAAVSQDNVPALQPGKKKKLR				ĺ			
LPPIKPRKQENSSQENENKTKVSGTKRIKHEIP KSKKEKENAPKPQDAAEVSSEQEKEQETLISQ KSIPEPLPAADMKKKIEGYQEFSAKPLASRVD PEKDNETDQGSNSEKVAEEAGEKGPTPPLPSA PLAPEKDSALVPGASKQPLTSPSALVDSKQES KLCCFTESPESEPQEASFPRLPHHTGHRWQTR MRRRMTNCPPWQITLPTAP 1041 2391 A 8646 113 1492 LLQEMCTKTIPVLWGCFLLWNLYVSSSQTIYP GIKARITQRALDYGVQAGMKMIEQMLKEKK LPDLSGSESLEFLKVDYVNYNFSNIKISAFSFP NTSLAFVPGVGIKAI.TNHGTANISTDWGFESP LFVLYNSFAEPMEKPILKNLNEMLCPIIASEVK ALNANLSTLEVLTKIDNYTLLDYSLISSPEITE NYLDLNLKGVPYPLENLTDPPFSPVPFVLPER SNSMLYIGIAEYFFKSASFAHFTAGVFNVTLS TEEISNHFVQNSQGLGNVLSRIAEIYILSQPFM VRIMATEPPIINLQPGNFTLDIPASIMMLTQPK NSIVETIVSMDFVASTSVGLVILGQRLVCSLS LNRFRLALPESNRSNIEVLRFENILSSILHFGVL PLANAKLQQGFPLPNPHKFLFVNSDIEVLEGF LLISTDLKYETSSKQQPSFHVWEGLNLISRQW RGKSAP 1042 2392 A 8672 538 170 ARRIARTRESKAAVSQDNVPALQPGKKKKLR							
KSIPEPLPAADMKKKIEGYQEFSAKPLASRVD PEKDNETDQGSNSEK VAEEAGEKGPTPPLPSA PLAPEKDSALVPGASKQPLTSPSALVDSKQES KLCCFTESPESEPQEASFPRLPHHTGHRWQTR MRRRMTNCPPWQITLPTAP 1041 2391 A 8646 113 1492 LLQEMCTKTIPVLWGCFLLWNLYVSSSQTIYP GIKARITQRALDYGVQAGMKMIEQMLKEKK LPDLSGSESLEFLKVDYVNYNFSNIKISAFSFP NTSLAFVPGVGIKALTNHGTANISTDWGFESP LFVLYNSFAEPMEKPILKNLNEMLCPHIASEVK ALNANLSTLEVLTKIDNYTLLDYSLISSPEITE NYLDLNLKGVFYPLENLTDPPFSPVPFVLPER SNSMLYIGIAEYFFKSASFAHFTAGVFNVTLS TEEISNHFVQNSQGLGNVLSRIAEIYILSQPFM VRIMATEPPIINLQPGNFTLDIPASIMMLTQPK NSIVETIVSMDFVASTSVGLVILGQRLVCSLS LNRFRLALPESNRSNIEVLRFENILSSILHFGVL PLANAKLQQGFPLPNPHKFLFVNSDIEVLEGF LLISTDLKYETSSKQQPSFHVWEGLNLISRQW RGKSAP 1042 2392 A 8672 538 170 ARRIARTRESKAAVSQDNVPALQPGKKKKLR							
KSIPEPLPAADMKKKIEGYQEFSAKPLASRVD PEKDNETDQGSNSEK VAEEAGEKGPTPPLPSA PLAPEKDSALVPGASKQPLTSPSALVDSKQES KLCCFTESPESEPQEASFPRLPHHTGHRWQTR MRRRMTNCPPWQITLPTAP 1041 2391 A 8646 113 1492 LLQEMCTKTIPVLWGCFLLWNLYVSSSQTIYP GIKARITQRALDYGVQAGMKMIEQMLKEKK LPDLSGSESLEFLKVDYVNYNFSNIKISAFSFP NTSLAFVPGVGIKALTNHGTANISTDWGFESP LFVLYNSFAEPMEKPILKNLNEMLCPHIASEVK ALNANLSTLEVLTKIDNYTLLDYSLISSPEITE NYLDLNLKGVFYPLENLTDPPFSPVPFVLPER SNSMLYIGIAEYFFKSASFAHFTAGVFNVTLS TEEISNHFVQNSQGLGNVLSRIAEIYILSQPFM VRIMATEPPIINLQPGNFTLDIPASIMMLTQPK NSIVETIVSMDFVASTSVGLVILGQRLVCSLS LNRFRLALPESNRSNIEVLRFENILSSILHFGVL PLANAKLQQGFPLPNPHKFLFVNSDIEVLEGF LLISTDLKYETSSKQQPSFHVWEGLNLISRQW RGKSAP 1042 2392 A 8672 538 170 ARRIARTRESKAAVSQDNVPALQPGKKKKLR	1						
PLAPEKDSALVPGASKQPLTSPSALVDSKQES KLCCFTESPESEPQEASFPRLPHHTGHRWQTR MRRRMTNCPPWQITLPTAP 1041 2391 A 8646 113 1492 LLQEMCTKTIPVLWGCFLLWNLYVSSSQTIYP GIKARITQRALDYGVQAGMKMIEQMLKEKK LPDLSGSESLEFLKVDYVNYNFSNIKISAFSFP NTSLAFVPGVGIKALTNHGTANISTDWGFESP LFVLYNSFAEPMEKPILKNLNEMLCPIIASEVK ALNANLSTLEVLTKIDNYTLLDYSLISSPEITE NYLDLNLKGVFYPLENLTDPPFSPVPFVLPER SNSMLYIGIAEYFFKSASFAHFTAGVFNVTLS TEEISNHFVQNSQGLGNVLSRIAEIYILSQPFM VRIMATEPPIINLQPGNFTLDIPASIMMLTQPK NSTVETIVSMDFVASTSVGLVILGQRLVCSLS LNRFRLALPESNRSNIEVLRFENILSSILHFGVL PLANAKLQQGFPLPNPHKFLFVNSDIEVLEGF LLISTDLKYETSSKQQPSFHVWEGLNLISRQW RGKSAP 1042 2392 A 8672 538 170 ARRIARTRESKAAVSQDNVPALQPGKKKKLR							KSIPEPLPAADMKKKIEGYQEFSAKPLASRVD
Continue			:				
MRRRMTNCPPWQITLPTAP 1041 2391 A 8646 113 1492 LLQEMCTKTIPVLWGCFLLWNLYVSSSQTIYP GIKARITQRALDYGVQAGMKMIEQMLKEKK LPDLSGSESLEFLKVDYVNYNFSNIKISAFSFP NTSLAFVPGVGIKALTNHGTANISTDWGFESP LFVLYNSFAEPMEKPILKNLNEMLCPIIASEVK ALNANLSTLEVLTKIDNYTLLDYSLISSPEITE NYLDLNLKGVFYPLENLTDPPFSPVPFVLPER SNSMLYIGIAEYFFKSASFAHFTAGVFNVTLS TEEISNHFVQNSQGLGNVLSRIAEIYILSQPFM VRIMATEPPIINLQPGNFTLDIPASIMMLTQPK NSTVETIVSMDFVASTSVGLVILGQRLVCSLS LNRFRLALPESNRSNIEVLRFENILSSILHFGVL PLANAKLQQGFPLPNPHKFLFVNSDIEVLEGF LLISTDLKYETSSKQQPSFHVWEGLNLISRQW RGKSAP 1042 2392 A 8672 538 170 ARRIARTRESKAAVSQDNVPALQPGKKKKLR			j		ł		
1041 2391 A 8646 113 1492 LLQEMCTKTIPVLWGCFLLWNLYVSSSQTIYP GIKARITQRALDYGVQAGMKMIEQMLKEKK LPDLSGSESLEFLKVDYVNYNFSNIKISAFSFP NTSLAFVPGVGIKALTNHGTANISTDWGFESP LFVLYNSFAEPMEKPILKNLNEMLCPIIASEVK ALNANLSTLEVLTKIDNYTLLDYSLISSPEITE NYLDLNLKGVFYPLENLTDPPFSPVPFVLPER SNSMLYIGIAEYFFKSASFAHFTAGVFNVTLS TEEISNHFVQNSQGLGNVLSRIAEIYILSQPFM VRIMATEPPIINLQPGNFTLDIPASIMMLTQPK NSTVETIVSMDFVASTSVGLVILGQRLVCSLS LNRFRLALPESNRSNIEVLRFENLSSILHFGVL PLANAKLQQGFPLPNPHKFLFVNSDIEVLEGF LLISTDLKYETSSKQQPSFHVWEGLNLISRQW RGKSAP							
GIKARITQRALDYGVQAGMKMIEQMLKEKK LPDLSGSESLEFLKVDYVNYNFSNIKISAFSFP NTSLAFVPGVGIKALTNHGTANISTDWGFESP LFVLYNSFAEPMEKPILKNLNEMLCPIIASEVK ALNANLSTLEVLTKIDNYTLLDYSLISSPEITE NYLDLNLKGVFYPLENLTDPPFSPVPFVLPER SNSMLYIGIAEYFFKSASFAHFTAGVFNVTLS TEEISNHFVQNSQGLGNVLSRIAEIYILSQPFM VRIMATEPPIINLQPGNFTLDIPASIMMLTQPK NSIVETIVSMDFVASTSVGLVILGGRLVCSLS LNRFRLALPESNRSNIEVLRFENILSSILHFGVL PLANAKLQQGFPLPNPHKFLFVNSDIEVLEGF LLISTDLKYETSSKQQPSFHVWEGLNLISRQW RGKSAP 1042 2392 A 8672 538 170 ARRIARTRESKAAVSQDNVPALQPGKKKKLR	1041	2391	A	8646	113	1492	I I OFMOTETTING WOOD I WAI VUCCOOTED
LPDLSGSESLEFLKVDYVNYNFSNIKISAFSFP NTSLAFVPGVGIKALTNHGTANISTDWGFESP LFVLYNSFAEPMEKPILKNLNEMLCPIIASEVK ALNANLSTLEVLTKIDNYTLLDYSLISSPEITE NYLDLNLKGVFYPLENLTDPFFSPVPFVLPER SNSMLYIGIAEYFFKSASFAHFTAGVFNVTLS TEEISNHFVQNSQGLGNVLSRIAEIYILSQPFM VRIMATEPPIINLQPGNFTLDIPASIMMLTQPK NSIVETIVSMDFVASTSVGLVILGQRLVCSLS LNRFRLALPESNRSNIEVLRFENILSSILHFGVL PLANAKLQQGFPLPNPHKFLFVNSDIEVLEGF LLISTDLKYETSSKQQPSFHVWEGLNLISRQW RGKSAP 1042 2392 A 8672 538 170 ARRIARTRESKAAVSQDNVPALQPGKKKKLR	'''		11	30-0	.13	1776	
NTSLAFVPGVGIKALTNHGTANISTDWGFESP LFVLYNSFAEPMEKPILKNLNEMLCPIIASEVK ALNANLSTLEVLTKIDNYTLLDYSLISSPEITE NYLDLNLKGVFYPLENLTDPPFSPVPFVLPER SNSMLYIGIAEYFFKSASFAHFTAGVFNVTLS TEEISNHFVQNSQGLGNVLSRIAEIYILSQPFM VRIMATEPPIINLQPGNFTLDIPASIMMLTQPK NSIVETIVSMDFVASTSVGLVILGGRLVCSLS LNRFRLALPESNRSNIEVLRFENILSSILHFGVL PLANAKLQQGFPLPNPHKFLFVNSDIEVLEGF LLISTDLKYETSSKQQPSFHVWEGLNLISRQW RGKSAP 1042 2392 A 8672 538 170 ARRIARTRESKAAVSQDNVPALQPGKKKKLR					ſ		
LFVLYNSFAEPMEKPILKNLNEMLCPIIASEVK ALNANLSTLEVLTKIDNYTLLDYSLISSPEITE NYLDLNLKGVFYPLENLTDPPFSPVPFVLPER SNSMLYIGIAEYFFKSASFAHFTAGVFNVTLS TEEISNHFVQNSQGLGNVLSRIAEIYILSQPFM VRIMATEPPIINLQPGNFTLDIPASIMMLTQPK NSIVETIVSMDFVASTSVGLVILGQRLVCSLS LNRFRLALPESNRSNIEVLRFENILSSILHFGVL PLANAKLQQGFPLPNPHKFLFVNSDIEVLEGF LLISTDLKYETSSKQQPSFHVWEGLNLISRQW RGKSAP 1042 2392 A 8672 538 170 ARRIARTRESKAAVSQDNVPALQPGKKKKLR							
ALNANLSTLEVLTKIDNYTLLDYSLISSPEITE NYLDLNLKGVFYPLENLTDPPFSPVPFVLPER SNSMLYIGIAEYFFKSASFAHFTAGVFNVTLS TEEISNHFVQNSQGLGNVLSRIAEIYILSQPFM VRIMATEPPIINLQPGNFTLDIPASIMMLTQPK NSTVETIVSMDFVASTSVGLVILGQRLVCSLS LNRFRLALPESNRSNIEVLRFENILSSILHFGVL PLANAKLQQGFPLPNPHKFLFVNSDIEVLEGF LLISTDLKYETSSKQQPSFHVWEGLNLISRQW RGKSAP 1042 2392 A 8672 538 170 ARRIARTRESKAAVSQDNVPALQPGKKKKLR	1			}	}	ł	
SNSMLYIGIAEYFFKSASFAHFTAGVFNVTLS TEEISNHFVQNSQGLGNVLSRIAEIYILSQPFM VRIMATEPPIINLQPGNFTLDIPASIMMLTQPK NSTVETIVSMDFVASTSVGLVILGQRLVCSLS LNRFRLALPESNRSNIEVLRFENILSSILHFGVL PLANAKLQQGFPLPNPHKFLFVNSDIEVLEGF LLISTDLKYETSSKQQPSFHVWEGLNLISRQW RGKSAP 1042 2392 A 8672 538 170 ARRIARTRESKAAVSQDNVPALQPGKKKKLR			i			l	
TEEISNHFVQNSQGLGNVLSRIAEIYILSQPFM VRIMATEPPIINLQPGNFTLDIPASIMMLTQPK NSTVETIVSMDFVASTSVGLVILGQRLVCSLS LNRFRLALPESNRSNIEVLRFENILSSILHFGVL PLANAKLQQGFPLPNPHKFLFVNSDIEVLEGF LLISTDLKYETSSKQQPSFHVWEGLNLISRQW RGKSAP 1042 2392 A 8672 538 170 ARRIARTRESKAAVSQDNVPALQPGKKKKLR					1		
VRIMATEPPINLQPGNFTLDIPASIMMLTQPK NSTVETIVSMDFVASTSVGLVILGQRLVCSLS LNRFRLALPESNRSNIEVLRFENILSSILHFGVL PLANAKLQQGFPLPNPHKFLFVNSDIEVLEGF LLISTDLKYETSSKQQPSFHVWEGLNLISRQW RGKSAP 1042 2392 A 8672 538 170 ARRIARTRESKAAVSQDNVPALQPGKKKKLR							
NSTVETIVSMDFVASTSVGLVILGQRLVCSLS LNRFRLALPESNRSNIEVLRFENILSSILHFGVL PLANAKLQQGFPLPNPHKFLFVNSDIEVLEGF LLISTDLKYETSSKQQPSFHVWEGLNLISRQW RGKSAP 1042 2392 A 8672 538 170 ARRIARTRESKAAVSQDNVPALQPGKKKKLR							
LNRFRLALPESNRSNIEVLRFENILSSILHFGVL PLANAKLQQGFPLPNPHKFLFVNSDIEVLEGF LLISTDLKYETSSKQQPSFHVWEGLNLISRQW RGKSAP 1042 2392 A 8672 538 170 ARRIARTRESKAAVSQDNVPALQPGKKKKLR					l		
PLANAKLQQGFPLPNPHKFLFVNSDIEVLEGF LLISTDLKYETSSKQQPSFHVWEGLNLISRQW RGKSAP 1042 2392 A 8672 538 170 ARRIARTRESKAAVSQDNVPALQPGKKKKLR	1			}	}	}	
LLISTDLK YETSSKQQPSFHVWEGLNLISRQW RGKSAP 1042 2392 A 8672 538 170 ARRIARTRESKAAVSQDNVPALQPGKKKKLR]	1		
1042 2392 A 8672 538 170 ARRIARTRESKAAVSQDNVPALQPGKKKKLR							
The state of the s	1042	2392	\overline{A}	8672	538	170	
				30,2	-30	1/0	LGGKKKKFKFFRLPKEFKKQLMYSPSNFKKM

	Tana	T > c	1055		T =	
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	i	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		!	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	ł	ł		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		İ		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ł		peptide		/=possible nucleotide deletion, \=possible
		ļ		sequence	}	nucleotide insertion
	i					TSLAGNTVQCLNKLKYVIYSAQYPAYGNITT
						LDMITSTDHVLEQDFWICFTFYSVKERQI
1043	2393	A	8688	359	17	GLKTRAPATPTFOREVLGPAKODMORRCPRI
10.0		[1 0000		l ''	GLMTSLLKPIKRRWRDYKRWKSGGFTGESC
]				1	HHADTLGDRGGLQGDHSELLQWQKRILRTE
			1		1	GEPSPKYISKNIFPICSYITGFL
1044	2394	A	8718	292	1490	
1044	2394	A	0/10	292	1490	GTVKTSVATPITAGHSCSSGGVLQVKSPATQS
	!		[[i	GFKFTSKMEDFNMESDSFEDFWKGEDLSNYS
	l .			1		YSSTLPPFLLDAAPCEPESLEINKYFVVIIYAL
	1	1	1			VFLLSLLGNSLVMLVILYSRVGRSVTDVYLL
	ļ	l	1	ł	l	NLALADLLFALTLPIWAASKVNGWIFGTFLC
	ļ					KVVSLLKEVNFYSGILLLACISVDRYLAIVHA
		ļ	İ			TRTLTQKRYLVKFICLSIWGLSLLLALPVLLFR
	1]	ļ	ļ	· .	RTVYSSNVSPACYEDMGNNTANWRMLLRIL
			1		ĺ	POSFGFIVPLLIMLFCYGFTLRTLFKAHMGOK
	1					HRAMRVIFAVVLIFLLCWLPYNLVLLADTLM
			1			RTQVIQETCERRNHIDRALDATEILGILHSCLN
		ĺ	1		İ	PLIYAFIGOKFRHGLLKILAIHGLISKDSLPKDS
		}			J	RPSFVGSSSGHTSTTL
1045	2395	A	8724	254	3184	FRANLAITVANRRGAQGGKMHTCCPPVTLEQ
1073	2393	^	6724	234	3104	
			İ			DLHRKMHSWMLQTLAFAVTSLVLSCAETIDY
		ľ	1			YGEICDNACPCEEKDGILTVSCENRGIISLSEIS
						PPRFPIYHLLLSGNLLNRLYPNEFVNYTGASIL
					Í	HLGSNVIQDIETGAFHGLRGLRRLHLNNNKL
		ł				ELLRDDTFLGLENLEYLQVDYNYISVIEPNAF
	!					GKLHLLQVLILNDNLLSSLPNNLFRFVPLTHL
						DLRGNRLKLLPYVGLLQHMDKVVELQLEEN
]			PWNCSCELISLKDWLDSISYSALVGDVVCETP
			1			FRLHGRDLDEVSKQELCPRRLISDYEMRPQTP
] .			LSTTGYLHTTPASVNSVATSSSAVYKPPLKPP
			1			KGTRQPNKPRVRPTSRQPSKDLGYSNYGPSIA
			1 :			YQTKSPVPLECPTACSCNLQISDLGLNVNCQE
						RKIESIAELQPKPYNPKKMYLTENYIAVVRRT
			·			DLLEATGLDLLHLGNNRISMIQDRAFGDLTN
				1		LRRLYLNGNRIERLSPELFYGLOSLOYLFLOY
			1			NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS
						1
						GVFSGLTLLRLNLRSNHFTSLPVSGVLDQLKS
			}			LIQIDLHDNPWDCTCDIVGMKLWVEQLKVG
)			VLVDEVICKAPKKFAETDMRSIKSELLCPDYS
						DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA
				. !		PASLGAGGGASSVPLSVLILSLLLVFIMSVFVA
						AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN
						MQYSVYGGGGGTGGHPHAHVHHRGPALPK
			, ,	J		VKTPAGHVYEYIPHPLGHMCKNPIYRSREGN j
					ĺ	SVEDYKDLHELKVTYSSNHHLQQQQQPPPPP
ļ				ļ		QQPQQQPPPQLQLQPGEEERRESHHLRSPAYS
				İ		VSTIEPREDLLSPVODADRFYRGILEPDKHCST
				ł	f	TPAGNSLPEYPKFPCSPAAYTFSPNYDLRRPH
				ļ	1	QYLHPGAGDSRLREPVLYSPPSAVFVEPNRNE
		'			l	YLELKAKLNVEPDYLEVLEKQTTFSQF
1046	2396	A	8736	28	452	
1070	2070	Λ.	0730	40	432	SPSAAGGLAWVSLALGSGSRGRDHSGSGVGT
	}					AMAGALVRKAADYVRSKDFRDYLMSTHFW
					İ	GPVANWGLPIAAINDMKKSPEIISGRMTFALC
						CYSLTFMRFAYKVQPRNWLLFACHATNEVA
			L			QLIQGGRLIKHEMTKTASA
1047	2397	A	8741	673	924	ALPGTPQQTVTLNTDGKVKSFTSPHSNPNLPP
İ				i		AKFFTSLQSLNWSSHLPPSPATESVGKRGNAK
						PPTTKLLHSSPLWNFFAQQL
1048	2398	A	8747	3	5054	PEVTKPSLSQPTAASPIGSSPSPPVNGGNNAKR
				1		

SEQ NO: nucl- cotid	of NO: of peptide	Met hod	SEQ ID NO: in USSN	Predicted beginning nucleotide location	Predicted end nucleotide location corresponding	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, l=Isoleucine, K=Lysine, L=Leucine,
uence	1		914 914	ng to first amino acid residue of peptide sequence	to last amino acid residue of peptide sequence	M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						VAVPNGQPPSAARYMPREVPPRFRCQQDHK VLLKRGQPPPPSCMLLGGGAGPPPCTAPGAN PNNAQVTGALLQSESGTAPDSTLGGAAASNY ANSTWGSGASSNNGTSPNPIHIWDKVIVDGS DMEEWPCIASKDTESSSENTTDNNSASNPGSE KSTLPGSTTSNKGKGSQCQSASSGNECNLGV WKSDPKAKSVQSSNSTTENNNGLGNWRNVS GQDRIGPGSGFSNFNPNSNPSAWPALVQEGTS RKGALETDNSNSSAQVSTVGQTSREQQSKME NAGVNFVVSGREQAQIHNTDGPKNGNTNSL NLSSPNPMENKGMPFGMGLGNTSRSTDAPSQ STGDRKTGSVGSWGAARGPSGTDTVSGQSNS GNNGNNGKEREDSWKGASVQKSTGSKNDS WDNNNRSTGGSWNFGPQDSNDNKWGEGNK MTSGVSQGEWKQPTGSDELKIGEWSGPNQPN SSTGAWDNQKGHPLLENQGNAQAPCWGRSS SSTGSEVEGQSTGSNHKAGSSDSHNSGRRSY
					·	RPTHPDCQAVLQTLLSRTDLDPRVLSNTGWG QTQIKQDTVWDIEEVPRPEGKSDKGTEGWES AATQTKNSGGWGDAPSQSNQMKSGWGELS ASTEWKDPKNTGGWNDYKNNNSSNWGGGR PDEKTPSSWNENPSKDQGWGGGRQPNQGWS SGKNGWGEEVDQTKNSNWESASKPVSGWG EGGQNEIGTWGNGGNASLASKGGWEDCKRS, PAWNETGRQPNSWNKQHQQQPPQQPPPPQ PEASGSWGGPPPPPPGNVRPSNSSWSSGPQPA TPKDEEPSGWEEPSPQSISRKMDIDDGTSAWG DPNSYNYKNVNLWDKNSQGGPAPREPNLPTP MTSKSASDSKSMQDGWGESDGPVTGARHPS WEEEEDGGVWNTTGSQGSASSHNSASWGQG GKKQMKCSLKGGNNDSWMNPLAKQFSNMG
1049	2399	A	8748	200	1387	LLSQTEDNPSSKMDLSVGSLSDKKFDVDKRA MNLGDFNDIMRKDRSGFRPPNSKDMGTTDS GPYFEKGGSHGLFGNSTAQSRGLHTPVQPLN SSPSLRAQVPPQFISPQVSASMLKQFPNSGLSP GLFNVGPQLSPQQIAMLSQLPQIPQFQLACQL LLQQQQQQLLQNQRKISQAVRQQQEQQLA RMVSALQQQQQQQRQPGMKHSPSHPVGPK PHLDNMVPNALNVGLPDLQTKGPIPGYGSGF SSGGMDYGMVGGKEAGTESRFKQWTSMME GLPSVATQEANMHKNGAIVAPGKTRGGSPY NQFDIIPGDTLGGHTGPAGDSWLPAKSPPTNK IGSKSSNASWPPEPQPGVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI
1049		А	8/48	200	1387	VPWKRQDEQLSLQVETLYLDSPAVIHLLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA PFALSALLYGANNNLVIYLQRYMDPSTYQVL

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO: in	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide seq- uence	seq- uence		USSN 09/496 914	location correspondi ng to first amino acid	to last amino acid residue of peptide	I=Isoleucine, K=Lysine, L=Leucine, M:=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
				residue of peptide sequence	sequence	Y=Threonine, V=Valine, W=Trypuphan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
v						SNLKIGSTAVLYCLCLRHRLSVRQGLALLLL MAAGACYAAGGLQVPGNTLPSPPPAAAASP MPLHITPLGLLLLILYCLISGLSSVYTELLMKR
						QRLPLALQNLFLYTFGVLLNLGLHAGGGSGP GLLEGFSGWAALVVLSQALNGLLMSAVMKH GSSITRLFVVSCSLVVNAVLSAVLLRLQLTAA
1050	2400	A	8758	3	1660	FFLATLLIGLAMRLYYGSR WVSSMGFEELLEQVGGFGPFQLRNVALLALP RVLLPLHFLLPIFLAAVPAHRCALPGAPANFS
						HQDVWLEAHLPREPDGTLSSCLRFAYPQALP NTTLGEERQSRGELEDEPATVPCSQGWEYDH SEFSSTIATESQWDLVCEQKGLNRAASTFFFA
				,		GVLVGAVAFGYLSDRFGRRRLLLVAYVSTLV LGLASAASVSYVMFAITRTLTGSALAGFTIIV MPLELEWLDVEHRTVAGVLSSTFWTGGVML
						LALVGYLIRDWRWLLLAVTLPCAPGILSLWW VPESARWLLTQGHVKEAHRYLLHCARLNGR
						PVCEDSFSQEAVSKVAAGERVVRRPSYLDLF RTPRLRHISLCCVVVWFGVNFSYYGLSLDVS GLGLNVYOTOLLFGAVELPSKLLVYLSVRYA
						GRRLTQAGTLLGTALAFGTRLLVSSDMKSWS TVLAVMGKAFSEAAFTTAYLFTSELYPTVLR
			-			QTGMGLTALVGRLGGSLAPLAALLDGVWLS LPKLTYGGIALLAAGTALLLPETRQAQLPETI QDVERKSAPTSLQEEEMPMKQVQN
1051	2401	Ā	8759	515	1625	EIRTPVAVSSAPSGDSEGDEEETTQDEVSSHTS EEDGGVVKVEKELENTEQPVGGNEVVEHEV TGNLNSDPLLELCQCPLCQLDCGSREQLIAHV
						YQHTAAVVSAKSYMCPVCGRALSSPGSLGR HLLIHSEDQRSNCAVCGARFTSHATFNSEKLP
						EVLNMESLPTVHNEGPSSAEGKDIAFSPPVYP AGILLVCNNCAAYRKLLEAQTPSVRKWALRR QNEPLEVRLQRLERERTAKKSRRDNETPEERE
						VRRMRDREAKRLQRMQETDEQRARRLQRDR EAMRLKRANETPEKRQARLIREREAKRLKRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF
1052	2402	Α	8763	1106	70	QLPVSGVELDSQLLGKMAFEEQNSSSLH RHGHGGRDRRGGGRVARPGGLGRYPGRGAA
						ASLVFVPTRRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI
 						ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP
						RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL
		ļ				LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY
1053	2403	A	8768	2	712	WLTFHLVR RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW
						PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK
						VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY
1054	2404	A	8769	344	527	SSLMKVENMSSNQDGNDSDEFM REATTLACRNSCWVFSRCSLGACKPTVCSMP
L	l					SLSRQGSQTLCLRLAEYCMESVDSQRLLLS

SEQ ID	SEO ID	Met	SEO	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in NO.	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine.
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine.
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	}	1	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	İ			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		1	}	peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
1055	2405	Α	8770	430	1104	QQESPAAGAARMNCKEGTDSSCGCRGNDEK
				1	'	KMLKCVVVGDGAVGKTCLLMSYANDAFPEE
				l		YVPTVFDHYAVTVTVGGKQHLLGLYDTAGQ
		1		!	1	EDYNQLRPLSYPNTDVFLICFSVVNPASYHNV
				!	İ	QEEWVPELKDCMPHVPYVLIGTQIDLRDDPK
		[1	1	TLARLLYMKEKPLTYEHGVKLAKAIGAQCYL
		İ	1			ECSALTQKGLKAVFDEAILTIFHPKKKKKRCS
<u></u>						EGHSCCSII
1056	2406	Α	8773	261	332	NPRIQLSGNSCCAGSCRVWLSEQ
1057	2407	A	8778	3	477	PAGIRHEQARGADRMGKCRGLRTARKLRSH
		1				RRDQKWHDKQYKKAHLGTALKANPFGGAS
		1				HAKGIVLEKVGVEAKQPNSAIRKCVRVQLIK
		Ì				NGKKITAFVPNDGCLNFÆENDEVLVAGFGR
	1	ŀ			ł	KGHAVGDIPGVRFKVVKVANVSLLALYKGK
		ļ				KERPRS
1058	2408	Α	8808	171	881	PGLSQEPSGSMETVVIVAIGVLATIFLASFAAL
		1				VLVCRQRYCRPRDLLQRYDSKPIVDLIGAME
ļ	}	}]		TQSEPSELELDDVVITNPHIEAILENEDWIEDA
						SGLMSHCIAILKICHTLTEKLVAMTMGSGAK
ļ				:		MKTSASVSDIIVVAKRISPRVDDVVKSMYPPL
						DPKLLDARTTALLLSVSHLVLVTRNACHLTG
1	1			}		GLDWIDQSLSAAEEHLEVLREAALASEPDKG
1059	2409	A	8809	246	757	LPGPEGFLQEQSAI
1039	2409	^	0007	246	757 .	MRLQGAIFVLLPHLGPILVWLFTRDHMSGWC
						EGPRMLSWCPFYKVLLLVQTAIYSVVGYASY
Į						LVWKDLGGGLGWPLALPLGLYAVQLTISWT VLVLFFTVHNPGLALLHLLLLYGLVVSTALI
}	1	l	1		}	WHPINKLAALLLLPYLAWLTVTSALTYHLWR
		1				DSLCPVHQPQPTEKSD
1060	2410	Ā	8810	304	381	PKLSVYPLQSHHCLSEPFQSLVCCLA
1061	2411	A	8820	1673	848	SCKTENLLEMWWFQQGLSFLPSALVIWTSAA
1	}	1		• •		FIFSYITAVTLHHIDPALPYISDTGTVAPEKCLF
	1	1	[GAMLNIAAVLCIATIYVRYKQVHALSPEENVI
						IKLNKAGLVLGILSCLGLSIVANFQKTTLFAA
	1					HVSGAVLTFGMGSLYMFVQTILSYQMQPKIH
1]	}		GKQVFWIRLLLVIWCGVSALSMLTCSSVLHS
1	1					GNFGTDLEQKLHWNPEDKGYVLHMITTAAE
1					.,	WSMSFSFFGFFLTYIRDFQKISLRVEANLHGL
<u></u>					*	TLYDTAPCPINNERTRLLSRDI
1062	2412	A	8824	1	763 ,	GGAPPASVPARESPVSGAQGSSRTRGHKRAA
1			[ĺ	GARAPQLCSSWQRRSAPAMSRGLQLLLLSCA
1						YSLAPATPEVKVACSEDVDLPCTAPWDPQVP
						YTVSWVKLLEGGEERMETPQEDHLRGQHYH
1						QKGQNGSFDAPNERPYSLKIRNTTSCNSGTYR
1	<u> </u>		(i		,	CTLQDPDGQRNLSGKVILRVTGCPAQRKEET
						FKKYRAEIVLLLALVIFYLTLIIFTCKFARLQSI
1					3	FPDFSKAGMERAFLPVTSPNKHLGLVTPHKT
1062	2412		9926	147	727	ELV CONTROL OF THE PROPERTY OF
1063	2413	Α'	8826	147	627	CETSTSSAGHAPCRHAAQGPPAEPTGLRLCSE
					, i	HQRLHAWPPGPRRPSLWPPKNGKWHSGKRT
				}		AGGRPQRRPSRRQSQRPSAWSGSPRMHSPGQ
]					KCSLMCPHRSQDSLSTAIFQRSPGANTGRALH
				j		CVLSKEMKSVQRSLGLSRIHLQSKRKIIHFVL
1064	2414	Ā	8835	2982	1860	TR
1004	2714	Λ.	0033	2702	1869	LKDTLKSQMTQEASDEAEDMKEAMNRMIDE
						LNKQVSELSQLYKEAQAELEDYRKRKSLEDV TAEYIHKAEHEKLMQLTNVSRAKAEDALSE
[.		MKSQYSKVLNELTQLKQLVDAQKENSVSITE
[Ì	ľ	HLQVITTLRTAAKEMEEKISNLKEHLASKEVE
	L					THE VITTER LAANEMICENSHINE THE ASKEVE

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion VAKLEKQLLEEKAAMTDAMVPRSSYEKLQS SLESEVSVLASKLKESVKEKEKVHSEVVQIRS EVSQVKREKENIQTLLKSKEQEVNELLQKFO
1065			2041			QAQEELAEMKRYSESSSKLEEDKDKKINEMS KEVTKLKEALNSLSQLSYSTSSSKRQSQQLEA LQQQVKQLQNQLAECKKQHQEVISVYRMHL LYAVQGQMDEDVQKVLKQILTMCKNQSQK K
1065	2415	A	8841	3	663	AAATAASLSPRGCRLRTPSSDVGPSRAPPPSA APLPTGRAQMSPSGRLCLLTIVGLILPTRGQTL KDTTSSSSADATIMDIQVPTRAPDAVYTELQP TSPTPTWPADETPQPQTQTQQLEGTDGPLVT DPETHKSTKAAHPTDDTTTLSERPSPSTDVQT DPQTLKPSGFHEDDPFFYDEHTLRKRGLLVA AVLFITGIIILTSGKCRQLSRLCRNHCR
1066	2416	A	8853	3806	2204	FVGEQEGGCEAGAGRGAQTYPGEAGERWFG RRRRGRVVSRKKMSLKSERRGIHVDQSDLL CKKGCGYYGNPAWQGFCSKCWREEYHKAR QKQIQEDWELAERLQREEEEAFASSQSSQGA QSLTFSKFEEKKTNEKTRKVITVKKFFSASSR VGSKKEIQEAKAPSPSINRQTSIETDRVSKEFIE FLKTFHKTGQEIYKQTKLFLEGMHYKRDLSIE EQSECAQDFYHNVAERMQTRGKVPPERVEKI MDQIEKYIMTRLYKYVFCPETTDDEKKDLAI QKRIRALRWVTPQMLCVPVNEDIPEVSDMVV KAITDIIEMDSKRVPRDKLACITKCSKHIFNAI KITKNEPASADDFLPTLIYIVLKGNPPRLQSNI QYITRFCNPSRLMTGEDGYYFTNLCCAVAFIE KLDAQSLNLSQEDFDRYMSGQTSPRKQEAES WSPDACLGVKQMYKNLDLLSQLNERQERIM NEAKKLEKDLIDWTDGIAREVQDIVEKYPLEI KPPNQPLAAIDSENVENDKLPPPLQPQVYAG
1067	2417	A	8855	1372	1513	SNMREVGCGWLVPVIPAFWEAEVGGSLEARS LRQAWATKQDPISKKK
1068	2418	Α	8856	1530	1583	PCRPGMECNSMISVHCNL
1069	2419	A	8857	1530	1583	PCRPGMECNSMISVHCNL
1070	2420	A	8866	293	1675 1675	PYPQGGYPQGPYPQEGYPQGPYPQGGYPQGP YPQSPFPPNPYGQPQVPFGQDPDSPQHGNYQ EEGPPSYYDNQDFPATNWDDKSIRQAFIRKVF LVLTLQLSVTLSTVSVFTFVAEVKGFVRENV WTYYVSYAVFFISLIVLSCCGDFRRKHPWNL VALSVLTASLSYMVGMIASFYNTEAVIMAVG IITAVCFTVVIFSMQTRYDFTSCMGVLLVSM VVLFIFAILCIFIRNRILEIVYASLGALLFTCFLA VDTQLLLGNKQLSLSPEEYVFAALNLYTDIINI FLYILTIIGRAKE*PSSSSLCPLRWHGWPGPCP WHGSASCTSPLSCPQAQPREKDASLQPSCMY TADTSIWTRCGHSMAPLVLPPPPRGTKATFPC HLLSTHCCMSPVCQPTPGTGGSTRSRGEGLSQ EVRVHVFPPVPAPQPGVEHPSPPPHPPGVLPS GDMRSGGLIPVLSPE ARGNTLYHLPRLCRKLNLRWFSASTLYDVOH
						DDKMGSNTFFKRNDCRYVMISCKADMAYDN VRHPFMI*SIKLIMEETYLNIIKAVYDRPTASII LNGEKLKVFPVRSGT*QGCSVWP
1072	2422	A	8870	33	658	MESVLSK YEDQITIFTDYLEEYPDTDELVWIL GKQHLLKTEKSKLLSDISARLWFTYRRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH LGRDWSWEKQKEQPKEYQRILQCFLDRKDC

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO: in	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide seq- uence	seq- uence		USSN 09/496 914	location correspondi ng to first	to last amino acid residue	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
				amino acid residue of peptide	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion CYSIHQMAQMGVGEGKSIGEWVLGPNTVAO
1073	2423	A	8879	146	412	GV*KNLA\LFDEW\NSLGLVYVSM\DNPSGSIA RFPKKLCRVLPL\SADTAGLTGP DFSV*GDVDIEVTCPICLOLLTEPLSLNCGLRL
				140	712	*QVCITA*IKESVIISGG*SSSPVCHTTFQPANL RTSRYLPT*SIKSLGPDEPQEG
1074	2424	A	8884	67	435	HLQGRSIRTLQLTGENEKNCEVSERIRRSGPW KEISFGDYICHTFQGDCWADRSPLHEAAAHG RLIALKTLIAQGVNVNLWTL/DRVSSLHEACL
1075	2425	A	8896	1294	248	*GPVACAKPYWKMVPRHGGTVTGPPLLMV RSGDRNGLTHQLGGLSQGSRNQSYRSRSRSR SRERPSAPRGIPFASASSSVYYGSYSRPYGSDK
					, ,	PWPSLLDKEREESLRQKRLSERERIGELGAPE VWGLSPKNPEPDSDEHTPVEDEEPKKSTTSAS
						TSEEEKKKKSSRSKERSKKRRKKKSSKRKHK KYSEDSDSDSDSETDSSDEDNKRRAKKAKK EKKKKHRSKKYKKKRSKKSRKESSDSSSKES
						QEEFLENPWKDRTKAEEPSDLIGPEAPKTLTS QDDKPLNYGHALLPGEGAAMAEYVKAGKRI
						PRRGEIGLTR*RNCHHLNAQVM**VVSRHRR MEAVRTAKREPESTVLMRREPLHPFNPRRET KERE
1076	2426	Α	8899	146	789	GRSTEAEKEPAFDERTGKGRRLPRAGEFHG*E *APGPGPRSFQVSRKMPEE\PPGARKHPFSGKS
						FYLDLPAGKNLQFLTGAIQQLGGVIEGFLSKE VSYIVSSRREVKAESSGKSHRGCPSPSPSEVR VETSAMVDPKGSHPRPSRKPVDSVPLSRGKE
1077	2427	A	8901	352	3	LLQKAIRNQK**CTVQQLSHCRLY\GEKTTAK RSQREHVQQQSQEHGKWPDLKGPR AKIGAYKYIQELWRKKQSDVMHFLLRVRCW
10//	2421	Α	8901	332	3	QYPALHRAGTEWQLSALHRAPRSTQPDKAC RLGYKAKQGYIIYRICVRRGGWKCPVPKAVT
10-2						\YGKPVHHGVN*LKFAQSLQSVAEEQ
1078	2428	A	8905	536	781	ACPAENREVPEMAAGQAPHAGPGAGPGQPA PALPFAATPGSRGQALCRGGRRRQHLHGPLH
1079	2429	A	8912	121	376 ·	RP*QAAPALHAGCQLAPHPPT NLIWKLCVTERRLVILDNYDLASE/YEANKYI CNRIIQFKPGQDKYFTLGLPTGSTPL*CYPKLI
						EYNKNGHLSFKYVKTFSMDEY
1080	2430	A	8920	381	1788	SSESPSDPGRMAMTWIVFSLWPLTVFMGHIG GHSLFSCEPITLRMCQDLPYNTTFMPNLLNHY
						DQQTAALAMEPFHPMVNLDCSRDFRPFLCAL YAPICMEYGRVTLPCRRLCQRAYSECSKLME
						MFGVPWPEDMECSRFPDCDEPYPRLVDLNLA GEPTEGAPVAVQRDYGFWCPRELKIDPDLGY SFLHVRDCSPPCPNMYFRREELSFARYFIGLIS
						IICLSATLFTFVTFLIDVTRFRYPERPIKCYAV WHMMVSLIFF\GFLLEDRVACNA\SIPAQYKA
			,			STVTQGSHNKACTMLFMILYFFTMAGSVWW VILTITWFLAAVPKWGSEAIEKKALLFHASA
						WGIPGTLTIILLAMNKIEGDNISGVCFVGLYD VDALRYFVLAPLCLYVVVGVSLLLAGIISLNR VDIEIDI *VENODEL VKEMIDIGVESIL VLADIL
						VRIEIPL*KENQDKLVKFMIRIGVFSILYLVPLL VVIGCYFYEQAYRGIWETTWIQERC
1081	2431	A	8922	56	420	EERTKMSTGPDVKATVGDISSDGNLNVAQEE CSRKGIVDEFFPLLSN*CIWTOPOGYPOSSYG
						TLANFVF\CSVRHGLALILQLCNFSIYTQQMN LSIAIPAMVNNTAPPSQPNASTERPST
1082	2432	Α	8923	355	1079	PFGTPSSTMAVVKNKCLMKGGKKGVKKKVV

	SEQ ID NO: of	SEQ ID NO: of	Met	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
USSN Ocasion Ogs/946 Ocasion Ogs/946 Ocasion Ogs/946 Ocasion Ogs/946 Ocasion Ogs/946 Ocasion Ogs/946 Ocasion Ogs/946 Ocasion		1		1			
United 1914	eotide	seq-		USSN	location	corresponding	
amino acid residue of peptide residue of peptide sequence Filtreonine, V=Valine, W=Tyotophan, Y=Tyosine, X=Ulaknow, Y=Sop codon, /=possible nucleotide deletion /=possible nucleotide deletion /=possible nucleotide deletion /=possible nucleotide deletion /=possible nucleotide deletion /=possible nucleotide deletion /=possible nucleotide deletion /=possible nucleotide deletion /=possible nucleotide deletion /=possible nucleotide deletion /=p	seq-	uence		I .		to last amino	
Persidue of peptide	uence	1		914			
Peptide							
	1					sequence	
GFFSKKDOYDVKAPAMPNIRNTGKTTLVART GGTQLABOGLAGLI-PEVSLADLO,DUSVARK FELITEDVODKNCL-TNFYGMDL-TCDKICSMV ERWSTMEARI-PUVKLTGOYPFILL-DCVGFTKK HNNQLKTSYA*HQQSRQIQKKMMEMT*EV DVFIRK VKAMLENPGFBERMELL-RGGOSS TJWPOPFIRSCPTRAVCOPYPLH DVFIRK VKAMLENPGFBERMELL-RGGOSS TJWPOPFIRSCPTRAVCOPYPLH DVFIRK VKAMLENPGFBERMELL-RGGOSS TJWPOPFIRSCPTRAVCOPYPLH DVFIRK VKAMLENPGFBERMELL-RGGOSS TJWPOPFIRSCPTANGENT-RCVGFTK GVAVSAGSPPIRKFERPCYGET GVALLESSISPIRCLSALAFANENP GVAVSAGSPPIRKFERPCYGET GVALLESSISPIRCLSALAFANENP GVAVSAGSPPIRKFERPCYGET GVALLESSISPIRCLSALAFANENP GVAVSAGSPPIRKFERPCYGET GVALLESSISPIRCLSALAFANENP GVAVSAGSPPIRKFERPCYGET GVALLESSISPIRCLSALAFANENP GVAVSAGSPPIRKFERPCYGET GVALLESSISPIRCLSALAFANENP GVAVSAGSPPIRKFERPCYGET GVALLESSISPIRCLSALAFANENP GVAVSAGSPPIRKFERPCYGET GVALLESSISPIRCLSALAFANENP GVAVSAGSPPIRKFWERT*GSPAK GVAVSAGSPPIRKFWERT*GSPAK GVAVSAGSPPIRKFWERT*GSPAK GVAVSAGSPPIRKFWERT*GSPAK GVAVSAGSPPIRKFWERT*GSPAK GVAVSAGSPPIRKFWERT*GSPAK GVAVSAGSPPIRKFWERT*GSPAK GVAVSAGSPPIRKFWERT*GSPAK GVAVSAGSPPIRKFWERT*GSPAK GVAVSAGSPRANENP GVAVS	1	i	ļ	ł		1	
QGTQLASDGLKGLLFEVSLADLQNDEVAFRK FELITEDYQDKNCLTNFYGMDLTCDKICSMV				· · · · · · · · · · · · · · · · · · ·	sequence		
FELITEDVQDKNCLTNFYGMDLTCDKICSMV ERWSTMEAHVDVKITTGYFFHLEVGFTIKK HNNQILKTSYA*HQQSRQIQKKMMEIMT*EV QTNDLKEVVNKLIPDIKGVTEK (VOFTYPLH DVFIRK VKMLENPGFERMELRGGOSS)			1		•	1	OGTOIASDGLKGLLFEVSLADLONDEVAERK
			1			}	
HNNQILKTSYA*HQGSRQICKKMEIMTER OTNDILEVIVALIPDNIGKDTEKVCPIYPLH							EKWSTMIEAHVDVKTTDGYFFHLFCVGFTKK
DVFIRKVKMLENPGFERNELRGGGSSS		İ					
1083		l	1				
1084	1002	2422		0040	20	205	
1084	1065	2433	A	8948	28	383	
1084							
1084	1	ļ]			
1085	1084	2434	Α	8950	156	318	
*TIYTSYDTAIPIS/GIYPKRMSSKCHQETCAR #TIYTSYDTAIPIS/GIYPKRMSSKCHQETCAR #FILAPFTATIKGQLTCPL/EERIDYMWYS HKYYIKVKRRL*YTITHYWNINILMFEIILW YSHKYY YSHYY YSHKYY YSH		İ		i			
MIPILAPPTATIKGKQLTCPLVEERIDYMWYS	1085	2435	A	8956	16	413	HMGQLGYFIQCWWECKRLISF\WKTI*QSPAK
HKYYIKVRRNL*VTITHYTWNLNILMFEIILW			ļ				
1086			1				
1086	[
1087 2437 A 8985 58 330	1086	2436	Δ	8962	868	1026	L
1087	1 1000	2430	Α.	0,02	800	1020	
**ERSVCAFHVCIQTYVCLQVYACMCVYYICM FYYSVYGCGLCTCVCMDVTICVCVQEFL	1087	2437	A	8985	58	330	
FYYSVYGGGLCTCVCMDYYICVCVQEFL							
1088	Ĺ			[[FVYSVYGCGLCTCVCMDVYICVCVQEFL
KYTVKRIKIHPTDLEKM.RNHLSDKD*YS/GV	1088	2438	Α	8989	394	404	
YKDLSKLNRRKTES* VKKWVKDLSRYFIKE VISMENKHKKIFSTS							
VISMENKHKKIFSTS							
1089 2439 A 8991 60 329 MALTPESPSSFPGLAATGSSVPEPPGGPNATL NSSWDSPTEPSSLEDLEATGTIGTLLSDMGVV GVEDNAYTLEVNSRYMRAVGIM*IHL	i		l				
NSSWDSPTEPSSLEDLEATGTIGTLLSDMGVV GVEDNAYTLEVNSRYMRAVGIM*HIL 1090 2440 A 8996 2 351 SNITITLT*MKKYDNTFCW*GCGQIGT/TLYC WQESKFIQAFWSKIQQYLA*ISIHILFDPAFLFL GGYPGGTQSVFLTGVLVSSVFYNMKMLHTR LLIAALFIIVQYWKQSKDHYI 1091 2441 A 8997 97 456 YPLPVCSYLSGPRGEHWNSLGKSSCPLPLPT LVSSRFKISKVIVVGDLSVGKTCLINR*GGAG AELGRVGPSLARWAGSRSQHLVPSQIVCKDS FDKNYKAPIGADFEMEFEVLGIPF 1092 2442 A 8999 548 811 SSFIKRHILIFEDDWHQTTCCHPPHHPV*RCQ FHIFYVSVQNSISPSLSVSSSHPDRPDHEVHQH RAAHIHIQHGQGPLGHGLVARVG 1093 2443 A 9002 3 2745 ALLGLQQPAQSLILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTTTAAGIKLIFFPDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGGETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEELIPLVVTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCTDSEPPRQEVPMCTDSEPPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMCTDSEPPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMCTDSEPRQEVPMCTDSEP	1089	2439	Ā	8991	60	329	
GVEDNAYTLEVNSRYMRAVGIM*IHL				0,,,	50	32)	NSSWDSPTEPSSLEDLEATGTIGTLLSDMGVV
WQESKFIQAFWSKIQQYLA*ISIHILFDPAFLFL GGYPGGTQSVFLTGVLVSSVFYNMKMLHTR LLIAALFIIVQYWKQSKDHYI 1091 2441 A 8997 97 456 YPLPVCSYLSGPRGEHWNSLGGKSSCPLPLPT LVSSRFKISKVIVVGDLSVGKTCLINR*GGAG AELGRVGPSLARWAGSRSQHLVPSQ\VCKDS FDKNYKAPIGADFEMERFEVLGIPF 1092 2442 A 8999 548 811 SSFIKRHILIFEDDWHQTTCCHHPHHPF*RCQ FHIFYVSVQNSISPSLSVSSSHPDRPDHEVHQH RAAHHHQHGQGPLGHGLVARVG 1093 2443 A 9002 3 2745 ALLGLQQPAQSLILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYPSALRDF VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHIGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGSEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMCTDSEPRQEVPMCTDSEP)	'		GVEDNAYTLEVNSRYMRAVGIM*IHL
GGYPGGTQSVFLTGVLVSSVFYNMKMLHTR LLIAALFIIVQYWKQSKDHYI 1091 2441 A 8997 97 456 YPLPVCSYLSGPRGEHWNSLGGKSSCPLPLPT LVSSRFKISK VIVVGDLSVGKTCLINR*GGAG AELGRVGPSLARWAGSRSQHLVPSQ\VCKDS FDKNYKAPIGAPFEMERFEVLGIPF 1092 2442 A 8999 548 811 SSFIKRHILIFEDDWHQTTCCHHPHHPF*RCQ FHIFYVSVQNSISPSLSVSSSHPDRPDHEVHQH RAAHIHHQHGQGPLGHGLVARVG 1093 2443 A 9002 3 2745 ALLGLQQPAQSLILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCTGDPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMCTDSEP	1090	2440	A	8996	2	351	SNITITLT*MKKYDNTFCW*GCGQIG/T/LIYC
LLIAALFIIVQYWKQSKDHYI 1091 2441 A 8997 97 456 YPLPVCSYLSGPRGEHWNSLGGKSSCPLPLPT LVSSRFKISKVIVVGDLSVGKTCLINR*GGAG AELGRVGPSLARWAGSRSQHLVPSQ\VCKDS FDKNYKAPIGADFEMERFEVLGIPF 1092 2442 A 8999 548 811 SSFIKRHILIFEDDWHQTTCCHHPHHPVF*RCQ FHIFYVSVQNSISPSLSVSSSHPDRPDHEVHQH RAAIIHHQHGQGPLGHGLVARVG 1093 2443 A 9002 3 2745 ALLGLQQPAQSLILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCHI/PEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTDSEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEPR RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY			,				
1091 2441 A 8997 97 456 YPLPVCSYLSGPRGEHWNSLGGKSSCPLPLPT LVSSRFKISKVIVVGDLSVGKTCLINR*GGAG AELGRVGPSLARWAGSRSQHLVPSQIVCKDS FDKNYKAPIGADFEMERFEVLGIPF 1092 2442 A 8999 548 811 SSFIKRHILIFEDDWHQTTCCHHPHHPPP*RCQ FHIFYVSVQNSISPSLSVSSSHPDRPDHEVHQH RAAHHHQHGQGPLGHGLVARVG 1093 2443 A 9002 3 2745 ALLGLQQPAQSLILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPLVVTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCTDSPEPRQEVPMCTGPEPRQEVP MCTGPERQEVPMCTDSEP RQEVPMCTDSEPRQEVPMCTDSEP RQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY							
LVSSRFKISKVIVVGDLSVGKTCLINR*GGAG AELGRVGPSLARWAGSRSQHLVPSQVCKDS FDKNYKAPIGADFEMERFEVLGIPF 1092 2442 A 8999 548 811 SSFIKRHILIFEDDWHQTTCCHHPHIPVF*RCQ FHIFYVSVQNSISPSLSVSSSHPDRPDHEVHQH RAAHHHQHGQGPLGHGLVARVG 1093 2443 A 9002 3 2745 ALLGLQPAQSLILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMCTDSEP	1001	2441	Δ	9007	07	450	
AELGRVGPSLARWAGSRSQHLVPSQVVCKDS FDKNYKAPIGADFEMERFEVLGIPF 1092 2442 A 8999 548 811 SSFIKRHILIFEDDWHQTTCCHHPHHPVF*RCQ FHIFYVSVQNSISPSLSVSSSHPDRPDHEVHQH RAAHIHQHGQGPLGHGLVARVG 1093 2443 A 9002 3 2745 ALLGLQQPAQSLILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLQGETLCSLQEADYEVASYGLQ HNCGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMCTDSEP	1071	2441	A	0991	97	430	
FDKNYKAPIGADFEMERFEVLGIPF 1092 2442 A 8999 548 811 SSFIKRHILIFEDDWHQTTCCHHPHHPF*RCQ FHIFYVSVQNSISPSLSVSSSHPDRPDHEVHQH RAAHHHQHGQGPLGHGLVARVG 1093 2443 A 9002 3 2745 ALLGLQQPAQSLILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTDSEP RQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY							
1092 2442 A 8999 548 811 SSFIKRHILIFEDDWHQTTCCHHPHHPF*RCQ FHIFYVSVQNSISPSLSVSSSHPDRPDHEVHQH RAAHHHQHGQGPLGHGLVARVG 1093 2443 A 9002 3 2745 ALLGLQQPAQSLILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTDSEP RQEVPMCTDSEP RQEVPMYTGSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMY							
FHIFYVSVQNSISPSLSVSSSHPDRPDHEVHQH RAAHIHQHGQGPLGHGLVARVG 1093 2443 A 9002 3 2745 ALLGLQQPAQSLILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY	1092	2442	A	8999	548	811	
A 9002 3 2745 ALLGLQQPAQSLILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTGSEPRQEVPMYTGPESPRQEVPMYTGSPRQUTGSPRQUTGTGSPRQUTGTGSPRQUTGTGSPRQUTGTGSPRQUTGTGSPRQUTGTGSPRQUTGTGSPRQUTGTGSPRQUTGTGSPRQUTGTGTGT							
TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMY					{		
AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LUTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMY	1093	2443	A	9002	3	2745	
VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LUTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMYT			l	ļ			
LKNNREISRIFHYIKSHKEQPĞRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFR YKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMYTGPESRQEVPMY				1			
VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMYTGPEPRQEVPMY				ļ	Í		
HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY	[{	ſ	ĺ		-
TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY							• • • •
PEGMFESFR YK CLSSYTSVKENFDKK GNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY		}			[
*RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY	[[[]			
RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY	-	ł	1				
GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY	ļ					İ	*RNGIISFTRT/INLIIGFSKNPKV**LWTNK*YP
PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY	}		1				
MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY	ľ	1	ľ	1	1		PESEBEADWCZDEEDBOEAD
RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY		ļ	l		ļ	ľ	MCTGPEAROEVPMCTDSFPROEVPMCTDSFP
						ļ	

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide		in in	nucleotide	location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		{	914	ng to first.	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	İ	1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	1	ļ	}	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		1	1	peptide	1	/=possible nucleotide deletion, \=possible
		<u> </u>	i	sequence	1	nucleotide insertion
						PICTOPISKQEDSMCTHAEINQKLPVATDFEFK
	ļ]]	J		LEALMCTNPEIKQEDPTNVGPEVKQQVTMVS
			1		[DTEILKVARTHHVQAESYLVYNIMSSGEIECS
			:		İ	NTLEDELDQALPSQAFIYRPIRQRVYSLLLED
		1	İ	ļ	1	CQDVTSTCLAVKEWFVYPGNPLRHPDLVRPL
	ĺ					QMTIPGGTPSLKILWLNQEPEIQVRRLDTLLA
				1		CFNLSSSREELQAVESPFQALCCLLIYLFVQV
					ł	DTLCLEDLHAFIAQALCLQGKSTSQLVNLQP
	ļ	1	İ	!		DYINPRAVQLGSLLVRGLTTLVLVNSACGFP
				i		WKTSDFMPWNVFDGKLFHQKYLQSEKGYA
		ĺ		ĺ		VEVL/CRTK*ISAHQIPQPEGSRLQGLHEGEQT
1094	2444	A	9021	97	834	HHWPSPLGLTPRREVGKTGLQLPQDGLWV
1034	2444	 ^	9021	97	834	AREACRAKTDFPGRRFRLWPSCCCRVIVGAE
	1	ļ	i			T*H\MAEPVSPLKHFVLAKKAITAIFDQLLEFV
						TEGSHFVEATYKNPELDRIATEDDLVEMQGY
	1	ļ			ļ	KDKLSIIGEVLSRRHMKVAFFGRTSSGKSSVI
		ļ				NAMLWDKVLPSGIGHITNCFLSVEGTDGDKA
						YLMTEGSDEKKSVKTVNQLAHALHMDKDLK AGCLVRVFWPKAKCALLRDDLVLVDGPGTD
						VTTELDSWIDKFCTKSSTREITNSGSDT
1095	2445	A	9022	1	537	LVLNSRVEDFVPPEGAGRTLPFALRPLAACW
				-	""	LLHRRARRSSALCPRPRSWGVSGGEGAGARE
						P*ITSSSCCLSAA/SHLSIQSPNMAGARRIRPQ
	1		<u> </u>			LAKEKIEGCHICTSVTPGEPQVFLGKDKAFTF
	1			,		DYVFDIDSQQEQIYIQCIEKLIEGCFEGYNATV
						FAYGQT\GAGKTYTMGTGFD
1096	2446	A	9029	1	285	FFFFNVCKSPKVPKPGCKEESTGTLFKNTLISL
						GQHSETPSLKKK\LAGYSGMCL*SQVLRRLRQ
1000						EDCLSPGGGNCRES*SCPYTPAWITERDPV
1097	2447	Α	9032	716	357	ARSTGFWGEILWCGFLKRSLALSPRVKCSGAI
	 -		-			LAHCNFRHAGFPPLSCLSLPNRWEYRRPPARP
	i		1			GKFFLVFLVETGFQC/G*DGLDLLTSRSACLG
1098	2448		0000			LPKCWDYRREPAASIIFQTTFFINSK
1098	2448	Λ	9038	230	652	KVVVMSCEDINISGSFYRNKLKYLAFLCKRTS
		İ		•		TNPSQGPYHLWVPSHIFWQTTCGRLPHKTKQ
	1			·		G*AALDHLKVFDRIPLPYDKKKQMAVSATLE
						VVRPKP*RKFAYLGHWAQKVDWKYQAMTA
1099	2449	Δ	9043	185	277	TMGEKRKVYYQKICYQKK
	2447	A	2043	707	372	IIFYSHQQCMRV/WQGCGDIETLIHCW*E*KÎI
	<u> </u>					HSL/WK/TV*QFLKRLYLHLPHNSVIAFLGISP
1100	2450	A	9045	763	584	RKIKTCPQNSCTSMLINAIHNDQKWKKINI
		*	7043	, 4,5	JU4	RQSLALSPRLECSGTISAHCRLCPLVFTPLSCL SLTSSWDYRRPPPHPANFLYFK*RRGF
1101 .	2451	Α	9050	275	2	LFFLRKVSNQFLSPSLLPVNFQGFVFAFLLLLL
	51	*	1030	2/3	-	FLL/FEMESLPVA/RVECSGTISAHCNLCLPGSS
•						DSPASAS*VAGITDMCRYTQLILFHAS
1102	2452	A	9053	449	1224	KTSMFWKFDLHSSSHIDTLLEREDVTLKELM
		- 1	1		1227	DEEDVLQECKAQNRKLIEFLLKAECLEDLVSF
	1					N*EEPPQDMDEKIRYKYPNISCELLTSDVSQM
			1			NDRLGEDESLLMKLYSFLLNDSPLNPLLASFF
	1		1	- 1		SKVLSILISRKPEQIVDFLKKKHDFVDLIIKHIG
			1	1		TSAIMDLLLRLLTCIEPPQPRQDVLN/WFKVQ
			1	ł	İ	RNL*HST*NVMDISKYVNLHWGLNKSHSLL*
		J	ļ			LLLQCVLQWLNEEKIIQRLVEIVHPSQEEDVS
		1	ľ	i	ļ	SLV
1103	2453	A	9058	403	3	GLHVYDFQVYREHILTLNVKKCSVSFWGLRE
T 103						,
1103		}	J	J		WLYLQMYEIIKSPRFPIIKMTDITKCW*GC\GA
1103						WLYLQMYEIIKSPRFPIIKMTDITKCW*GC\GA AGMQI/H/CW\WCVNVGKFWEMS*YYLLKLSI ST/PYDPAIPLLGIYL*ETRVYIHPKTCMRMLIA

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl- eotide seq- uence	peptide seq- uence		in USSN 09/496 914	nucleotide location correspondi ng to first amino acid	location corresponding to last amino acid residue of peptide	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
				residue of peptide sequence	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion APFVLAVNC
1104	2454	A	9064	75	393	KWLFSSLNITGRGDIIGHLKWLDCR\NCSSFPI KRNRQTHSTESNKLKAGHSFGYN*LIH*NS\V KTDCGCGANSKGVVVVMKV\KTAQQKQTTS YMQIGTTKNSRAT
1105	2455	A	9065	366	778	DLLILRNLAFPELKRRNCISRFYLAYHLHKIYS RSILLCNNCSGFYILSL*QYDVFFFNYFFFRDR AWPCCPGWSAAWLTIVILAHYRRPGLERSCC LSLSSSWDHRRVPPCPANF*/YFSMGFTAFPRL VLNS*TQGI
1106	2456	Α	9083	673	816	ESGSLIH*WWENKPAQPLWWEI*QHVQKLPT HFPCDPAIPLLGICPED
1107	2457	Α	9086	580	18	KPSSGSFIRAIYIFLSTAHVPALFSVLVRTKLT* AFSQSSVLWAHKQQKTSLSLVIR/ERLQIKTA VRENFLPIRLAKILKLDNVKCWQG/SGSNMSL I/HCWWEYNVIHIIWNSVTFPRKVEHVYITYA PEISVR*IHGGLPTLVHQETHTSVFRGAPSVIP ETR\CRPTKESINKLLHIYTMEHYGDENK
1108	2458	A	9093	540		GGNDCSVTPTTEPGRKEIT*KRKF*EKTDRLP GA/PPSRTPPTPYPCPHGDRLLPPSRPLPAGPA SAFPPAERSRGHRRASL*RARWSAAVPRRSA GSASEPVQSRWLRLPVGSDSPPAVPVRVCPAP DSRPAAPGSRLPDPGLDSPAPSRTPSSSVD*GG QRPPPPSGDSLSPPGCCRY
1109	2459	A	9099	1255	1425	HESYHVNPNLCNPVAPTSGAHSIG*KWPSWL GAVAHSCNPSTLVGRGGRITRGQELR
1110	2460	A	9103	242	70	EEQFFFFAVGMFP*VDFLAPASGELWDRLRLT CSRPFTRHQSFGLAFLRVCSSLDSLDDSVVGP SALLSSVL/NQGGRNVLEAREAAKHPTI*RQS LLRKQRNKRMAIP
1111	2461	Α	9110	189	121	SFLSVRLECNGAIMAHCALPLPG
1112	2462	A	9113	100	910	RRRGGGSRPRRTPVPAPGPGPSFGMDVRFYP AAAGDPASLDFAQCLGYYGYSKFGNNNNYM NMAEANNAFFAASEQTFHTPSLGDEEFEIPPIT PPPESDPALGMPDVLLPFQALSDPLPSQGSEFT PQFPPQSLDLPSITISRNLVEQDGVLHSSGLHM DQSHTQVSQYRQDPSLIMRPSST*PDAARSG VMPPAQLTTINQSQLSAQLGLNLGGASMPHT SPSPPASKSATPSPSSSINEEDADEANRAIGEK RAAPDSGKKPKTPKK
1113	2403	A	9120	3452	3051	FLRPSFALVPQAGVQWCALSWLQPPSPRFK*F SCLSLPSSWDYRHVPPRPANFFVLLVETGFLH VGQAGHEPLTSGDPPASASQSAGITGVSHQA WPSFFIFSRDTVLLCCSGWSRTSGLKQSACLS LLKCWDY
1114	2464	A	9122	152	377	NQLPLQQWTFFIYETGFCSVAQAGVQCRDHS SLHP*PPG\SSDPPAPPS*VLGITGQRYHACLII YLYVQTVPQRV
1115	2465	A	9124	553	981	QRPLLRQQLGSWPTCRSLEGDLASPW**RLPG SPRMRRSGT/ATLNLPLSPQGTVRTAVEFQVM TQTQSLSFLLGSSASLDCGFSMAPGLDLISVE WRLQHKGRGRGDLHLPDHHLSVPSSADHPA QQPSQFNGRNLYFLPLFR
1116	2466	A	9135	48 .	410	SASHEPAEHDGGADSLSASQPPRPAGRPAGA QHVHVPPWTDVLAGQDRRAPTAGDGAPWP APGGHVPSTRPHDPAEFHADEAAGRGGRGLQ PAAPHALPAGLPHGPPAPA/PAEGGGTP*GSA GAGGP*GSPAGRACGAAGCRPRPPRPAASSA *NSAGS*GLVEGT*PPGAGHGAPSPAVGARLS

	S ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO:		NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nuc	. 1	peptide	ł	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotic		seq-	ł	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-		uence	 -	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uene	te			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
.[amino acid residue of	of peptide	T=Threonine, V=Valine, W=Tryptophan,
					peptide	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
					sequence		/=possible nucleotide deletion, \=possible nucleotide insertion
\vdash	~		ļ		sequence	 	CPARTSVOCCTUTOTARACORACIO CONTEAR
	İ						CPARTSVQGGTWTC*APAGRPAGLGGWEAE RESAPPSCSAGS*DAD*GAEPWGAGSRSWGS
1111	7	2467	A	9141	380	939	KSGHWAKECLQPRIPPRPCPICVGPHWKSDCP
			1	71,11	300	/3/	TCPGAVPRAPGTLPQGSLTDSFPDLLSLVAED
			l			1	*CCLMASEASWTIT\ELWVTLTVEGKSVP/CL
						}	NTEATHSTLPSFQGPVSLASITVVGIDGQASKP
1	- 1		}	ļ			LKTPQLWCQLGQYSFMHYFLVIPTCPVPLLG*
L			l	1		•	GILTKLSAFLTIPRLQPHLIAALSPSS
1118	8	2468	A	9154	471	2	AAGQVVVEVTSHLYLCITSDAAGLRLLPPAES
1							ERGEGGHCPAEAPLPPRPQYCLAKHPLLRKLP
1				!		}	EEKIKLDPYLTQHTKINSKQIKYLS/VRAKTTQ
	İ			į		į	LVEGNIGVNLQNTELKQH*INGFLDTTPEAQE
L							TKEKTNKLNFIKKVKRQLAEWEKIFQIA
1119	9	2469	Α	9155	2	3187	ACPRLARRRRRVRSLRRRRGWLRARWSRGO
	1			[NNMAARRITQETFDAVLQEKAKRYHMDASĞ
1							EAVSETLQFKAQDLLRAVPRSRAEMYDDVHS
1	ŀ					}	DGRYSLSGSVAHSRDAGRESLRSDVFSGPSFR
l							SSNPSISDDSYFRKECGRDLEFSHSNSRDQVIG
-							HRKLGHFRSQDWKFALRGSWEQDFGHPVSQ
l							ESSWSQEYSFGPSAVLGDFGSSRLIEKECLEK
1		Ì		j		•	ESRDYDVDHPGEADSV/LRGGSQVQARGRAL
1							NIVDQEGSLLGKGETQGLLTAKGGVGKLVTL
			ľ				RNVSTKKIPTVNRITPKTQGTNQIQKNTPSPD
	1	J					VTLGTNPGTEDIQFPIQKIPLGLDLKNLRLPRR KMSFDIIDKSDVFSRFGIEIIKWAGFHTIKDDIK
-							FSQLFQTLFELETETCAKMLASFKCSLKPEHR
							DFCFFTIKFLKHSALKTPRVDNEFLNMLLDKG
							AVKTKNCFFEIIKPFDKYIMRLQDRLLKSVTP
1	ł						LLMACNAYELSVKMKTLSNPLDLALALETTN
							SLCRKSLALLGQTFSLASSFRQEKIL*AVGLQ
1]						DIAPSPAAFPNFEDSTLFGREYIDHLKAWLVS
							SGCPLQVKKAEPEPMREEEKMIPPTKPEIOAK
	İ						APSSLSDAVPQRADHRVVGTIDQLVKRVIEGS
1	i						LSPKERTLLKEDPAYWFLSDENSLEYKYYKL
ł	- 1						KLAEMQRMSENLRGADQKPTSADCAVRAML
							YSRAVRNLKKKLLP\WQRRGLLRAQG\LRG\
1		ĺ					WKARRA\TTGTQTLLFLRAPGLKHHGRQAPG
		l				,	LSQAKPSLPDRNDAAKDCPPDPVGPSPQDPSL
1	i	İ			ľ	' <u> </u>	EASGPSPKPAGVDISEAPQTSSPCPSADIDMKT
						ı	METAEKLARFVAQVGPEIEQFSIENSTDNPDL
1		İ		1			WFLHDQNSSAFKFYRKKVFELCPSICFTSSPH
1		- 1		ļ	1		NLHTGGGDTTGSQESPVDLMEGEAEFEDEPP
		ſ	[ſ	ĺ	PREAELESPEVMPEEEDEDDEDGGEEAPAPG
1	1	Ì	1				GAGKSEGSTPADGLPGEAAEDDLAGAPALSQ ASSGTCFPRKRISSKSLKVGMIPAPKRVCLIQE
1	- 1	ł	Ì	1	ł	1	PKGECPPVGTVASSTVLGWWAVRVRRDRWR
							HFNPKEFCAPLQNVSRHSCFPVV
1120		2470	A	9163	124	207	PPRACRPCPRACPCPPT*KCSQPVSWPC
1121		2471	A	9166	272	523	PMSSLQGCFYTFKCIIFKGIFLLISNI.IAF**EK
			[- 1	1		V/CSHITDSLKFIGKGWVGMVTHACNPGTLG
L					1	[G*GGWIA*VREFETSLGNM
1122		2472	C	9170	442	236	MNRRFLRPADCHSGMRGTENGACSEGESQI
1			J		i		HCGAGGEGVQLVHVVNQPENGCLQFDSTHIT
		ļ	- 1	- 1			FSKRQN*
1123	1:	2473	A	9171	10	423	MVDRSPLLTSVIIFYLAIGAAIFEVLEEPHWKE
1	- 1	1	- 1	1	1		AKKNYYTQKLHLLKEFPCLGQEGLDKILEVV
1			ļ	1	İ	1	SDAAGQGVAITGNQTFNNWNWPNAMIFAAT
l		İ	1	- 1	ļ	1	VITTIGYGNVASKTPGGRLFCGFYGLFGVPFC
L						1	LTWINALGKFFG
-							,

CEOTO	CEO ID	Met	Loro	D., 45-4-1	I B 12 4 1 1	
SEQ ID NO: of	SEQ ID NO: of	hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		İ		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
į		1	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		Ì		peptide		/=possible nucleotide deletion, \=possible
			<u> </u>	sequence		nucleotide insertion
1124	2474	Α	9173	3	374	GPSPSLLVLLPQEPGGTGTPVRAGAGAGMWL
				!		WEDQGGLLGPFSFLMLMLLLETRNPVNACLL
						TGSLFVLLGVFSFEPVPSCRALQELKPRDRISA
1125	0.475	L	0170			IAHRGGRHDPPENTLGAIR/QGS**WSNRR
1123	2475	A	9179	704	188	ESSSGLLFQCFQGIHVQKLTLQARPTLFSWWL
				ł		CSKPPKETGELENAESGGDGGRRGGKQDNV
						AWWRRM\QKG\DFPWDDEDFPQSGPFGGQA
		1			'	LPMGFFYLYFRDPGREITWKHFVQYYLARGL VDRLEVVNKQSVRVIPAPGTSSEVRGEFKAE
Í					[YCRHKFISCKNVVFYFFQ
1126	2476	A	9183	153	233	MEYMAESTDRSPGHILCCECGVPISPN
1127	2477	A	9185	1	321	LTGQLGSILLRVFSKSRAGLGARKLKAYRTM
****	2477	1 "	1103	}	321	EYMAESTDRSPGHILCCECGVPISPNPAOY\CV
ľ			1	1		ACLRSSFHIYHCIPKLFIHPFSKTSSSAFITPSHY
						LTFFSTIS
1128	2478	A	9186	183	847	VLKFLLLQTMDEQSQGMQGPPVPQFQPQKAL
						RPDMGYNTLANFRIEKKIGRGQ\FSEVYRAAC
		Į				L\LDGVPVALKKVQIFDLMDAKARADCIKEID
						LLKQLNHPNVIKYYASFIEDNELNIVLELADA
						GDLSRMIKHFKKQKRLIPERTVWKYFVQLCS
						ALEHMHSRRVMHRDIKPANVFITATGVVKLG
						DLGLGRFFSSKTTAAHSLVGTPYYMSPERIHD
						NG
1129	2479	A	9190	1	370	GTSWKIPSAAVSESSPNGAAYASGLPCGVRG
						PPWAGLALLPSPTLMALLRRPTVSSDLDNIDT
						RATTIKIRVVATITRARIEDMRHSATALTRPD
1130	2480		9194	131	487	ATTAQIPKLPVTTVCNRRANPGIPPSVL
1130	2400	А	9194	131	48/	AYLKRLPVPESITGFARLTVSEWLRLLPFLGV
			ļļ	•		LALLGYLAVRPFLPKKKQQKDSLINLKIQKEN PKVVNEINIEDLCLTKAAYCRCWRSKTFPAC
						DGSHNKHNELTGDNVGPLILKKKE
1131	2481	A	9201	184	605	KELVDEKSERGRAMDPVSOLASAGTFRVLKE
			/201	10.	005	PLAFLRALELLFAIFAFATCGGYSGGLRLSVD
			1	ĺ		CVNKTESNLSIDIAFAYPFRLHQVTFEG\PTCE
			1			GKERHKLALIGDSSSSAEFFGTVAGFAFLYSL
		•	l i			AATGVYIFFQNKY
1132	2482	A	9206	1	852	GGGRAGAGSRDMGSTDSKLNFRKAVIQLTTK
						TQPVEATDDAFWDQFWADTATSVQDVFALV
						PAAEIRAVREESPSNLATLCYKAVEKLVQGA
			l i			ESGCHSEKEKQIVLNCSRLLTRVLPYIFEDPD
	ĺ					WRGFFWSTVPGAGRGGQGEEDDEHARPLAE
l						SLLLAIADLLFCPDFTVQSHRRSTVDSAEDVH
l	ļ			İ		SLDSCEYIWEAGVGFAHSPQPNYIHDMNRME
ļ	}		,	j	J	LLKLLLTCFSEAMYLPPAPESWQH/RTHWFSS
	ì					FVSSENRHALPLFTSLLNTVCAYDPVEYGIPY
1133	2483	A	9208	1165	1463	NHLY
1133	2703	л	1200	1105	1403	GPRARVQGFSGADIVKFMALGSMYLVLTLIV AKVLRGAEPCCGPLKNRVLRPCPLP/VPLPPP
	- 1					HPOPSRGNPVGCLPTYKVVYKLLSWPLHSNS
1	ļ					NVYFIV
1134	2484	A	9210	66	1586	MAGAGPKRRALSAPVAEEKEEAREKIMAAK
					-500	RADGAAPAGEGEGVTLQGNITLLKGVAVIVV
	ļ	ĺ		1		AIMGSGIFVTPTGVLKEAGSPGLALVVWAAC
				1		GVFSIVGALCYAELGTTISKSGGDYAYMLDV
	İ			1		YGSLPAFLKLWIELLIIRPSSQYIVALVFATYL
}	ł			1		LKPLFPTCPVPEEAAKLVACLCVLLLTAVNC
		ļ			1	YSVKAATRVQDAFAAAKLLALALIILLGFVQI
1	1			l	l	GKGDVSNLDPNFSFEGTKLDVGNIVLALYSG
Į.	1	1		!	ſ	LFAY.GGWNYLNFVTEEMINPYRNLPLAIIISLP

NOT of much could be provided to the composition of	SEQ ID	SEQ ID	Met	SEQ	Predicted	Dendicted and	I A - i - a id - a id - i - a id - a id - i - a id - a i
Deptide					P .	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Sequence USSN Ocarison Distant Dista		l .	1104	1			D=Aspartic Acid, E=Glutamic Acid,
1135	1		ŀ		1		r=Pnenylalanine, G=Glycine, H=Histidine,
1136							1=isoleucine, K=Lysine, L=Leucine,
mino acid residue of peptide residue of peptide sequence purposite s		daice		1		1	O-Clutomine, N-Asparagine, P-Profine,
Pepside of peptide sequence	""""	İ		714			T-Thronging V-V-V-V-V-V-V-V-V-V-V-V-V-V-V-V-V-V-V-
Popsible sequence	1		ŀ				V=Turcsing V=Valine, w=Tryptopnan,
	1				l .	sequence	/=nonsible publication deletion \-nonsible
INTL-VYVLTILA-YFTILSTEGMLSSEAVAUPE	1				,	<u> </u>	
GNYHLGVMSWIIPVFVGLSCFGSVNGSLFTSS RLFPVGSRGGHLPSILSMIHPQLLTPVFLAFT CVMTLFYAFSKDIFSVINFFSFFWLCVALAII GMIWLERIKPELERIKFWALALPPFLACLF LLAVSFWKTTPWSVASDFTILISGLPVYFFGW WWKNKFWAPPGHLSPRPSCVRSSGWVPQ WWKNKFWAPPGHLSPRPSCVRSSGWVPQ WWKNKFWAPPGHLSPRPSCVRSSGWVPQ WWKNKFWAPPGHLSPRPSCVRSSGWVPQ WWKNKFWAPPGHLSPRPSCVRSSGWVPQ MQETLRNLASIGEKWKDQNLEDQYKNFRNIL SLIGERVFDVANFTGEWSLLDPSOKALYREV MQETLRNLASIGEKWKDQNLEDQYKNFRNIL SLIGERVFDVANFTGEWSLLDPSOKALYREV MQETLRNLASIGEKWKDQNLEDQYKNFRNIL SLIGERVFDVANFTGEWSLLDPSOKALYREV MQETLRNLASIGEKWKDQNLEDQYKNFRNIL SLIGERVFDFLAVSMFPHFARFG DLVFAKNKGYPHVPARDDLADGAVKPPN KYPIFFGTHETAFLGFDCLAFYPDKCKDKYGK PPKKKGPNEGLWEIQNNPHASYSAPPYSSSD SEAPLANPAGSDADEDDEGRGVMAYTAVT ATAASDEMESDSDSDSSDSSDSSGLKRKTPALK MVSKRARKASSDLOQASVSPSEEDRSSSSS SSSSDSDVSVKKPPRGKPAEFPLPKPGRK APASDSDSKADSDGAAFPPVAMARSSSSS SSSSSDSDVSVKKPPRGKPAEFPLPKPGRK APASDSDSKADSDGAAFPPVAMARSSSSS SSSSSDSDVSVKKPPRGKPAEFPLPKPGRK APASDSDSKADSDGAAFPPVAMARSSSSSS SSSSSDSDVSVKKPPRGKPAEFPLPKPGRK APASDSDSKADSDGAAFPPVAMARSSSSS SSSSSDSDVSVKPPRGKPAEFPLPKPGRK APASDSDSKADSDGAAFPPVAMARSSSSSS SSSSSDSDVSVKPPRGKPAEFPLPKPGRK APASDSDSKADSDGAAFPPVAMARSSSSSS SSSSSDSDVSVKPPRGKPAEFPLPKPGRK APASDSDSKADSDGAAFPPVAMARSSSSSS SSSSSDSDVSVKPPRGKPAEFPLPKPGRK APASDSDSKADSDGAAFPPVAMARSSSSSS SSSSSDSDVSVKPPRGKPAEFPLPKPGRK APASDSDSKADSDGAAFPPVAMARSSSSSS SSSSSDSDVSVKPPRGKPAEFPLPKPGRK APASDSDSKADSDGAAFPPVAMARSSSSSS SSSSSDSDVSVKPPRGKPAEFPLPKPGRK APASDSDSKADSDGAAFPPVAMARSSSSSS SSSSSDSDVSVKPPRGKPAEFPLPKPGRK APASDSDSKADSDGAAFPPVAMARSSSSSS SSSSSDSDVSVKPPRGKPAEFPLPKPGRK APASDSDSKADSDGAAFPPVAMARSSSSSS SSSSSDSDVSVKPPRGKPAEFPLFVRGRK APASDSDSKADSDGAAFPPVAMARSSSSSS SSSSSDSDVSVKPPRGKPAEFPLFLERPFGCAF TPTPQCKPETTFLDQCCSSPVLDCPHRNCNK APASDSDSKADSDGAAFPPVAMARSSSSSS SSSSSDSDVSVKPPRGKPAEFPLFLARPFGLIARP TTPQGKPETTFLDQCCSSPVLDCPHRNCNK APASDSADAFPAEFPLIARP TTPQGKPETTFLDQCCSSPVLDCAPHA SPGAGPTSTDLARPGGRGAKATTRSRSTDKR ARABAAPGELARPAEPTPPQLIA BPTATFTTTTTTTTCTTCTSTTDLAFT TTPQGKPETTFLDQCCSSPVLDCAPHA SPGAGPTSTDLAFT TTPQGKPETTFLDQCCSSPVLDCAPHA SPGAGPTSTDLAFT TTPQGKPETTFLDQCSSPSTANA NAGREPAAAAPGEDLAFT TTPQGKPETTFLDQCSSPS	<u> </u>	 	 		sequence		
RIFFVGSREGHLPSILSMIHPQLLTPYPSILYTI CVMTLYASKDIPSVINFSSFRWICVALAII GMWURHEKPELERPIKVNILAIPYFFILACLE LIAVSYEWITTPWSVASDFTILISGLPVYFFOV WWKNKEKWAPPGHLSPRPSCVRSSCMVVPQ WWKNKEKWAPPGHLSPRPSCVRSSCMVVPQ WWKNKEKWAPPGHLSPRPSCVRSSCMVVPQ WWKNKEKWAPPGHLSPRPSCVRSSCMVVPQ WWKNKEKWAPPGHLSPRPSCVRSSCMVVPQ WWKNKEKWAPPGHLSPRPSCVRSSCMVVPQ WWKNKEKWAPPGHLSPRPSCVRSSCMVVPQ WWKNKEKWAPPGHLSPRPSCVRSSCMVVPQ WWKNKEKWAPPGHLSPRPSCVRSSCMVVPQ WWKNKEKWAPPGHLSPRPSCVRSSCMVVPQ WWKNKEKWAPPGHLSPRPSCVRSSCMVVPQ WWKNKEKWAPPGHLSPRPSCVRSSCMVPQ WWKNKEKWAPPGHLDSPGVRSKNI YREV MOETLARILASIGEKWKDQNIEDQVKNIPRNIN RSLIGERVDENTERHICGTSSQIPDTLNK RSRRRSSRYBRCSRFPPGPLAVSMPPHAFKPG WYPFFFGTHETAFLGPRDLFPYDKCKDKYGK KYPFFFGTHETAFLGPRDLFPYDKCKDKYGK APSASDSDSKAADSDGAKPPPVNAKPSPVSSD SEAPEANPADGSDADEDBEGRGWMAYTAVT ATAASDBMESDASDSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS	l	}	}				
CVMTLFY_AFSKDIFSVINFSFFNWLCVALACE GMIW_RHRRPELEDERPKYNLL_DEST_CLACE LIAVSFWKTTPWSVASDFTILLSCLPVYFPCIC LIAVSFWKTTPWSVASDFTILLSCLPVYFPCIC WWKNKPK_WAPPGILSPRRSCVS_VSS_CMVVPQ WWKNKPK_WAPPGILSPRRSCVS_VSS_CMVVPQ WWKNKPK_WAPPGILSPRRSCVS_VSS_CMVVPQ DLVAFEDVAVWFTOFE_WSLLDPSQKNLYREV MQETLRNI_ASIGEK_WKDQNIEDQ_YKNPRNNL RSLLGEK_VMDNTEDH_CGETSQLPDFTTM. RSLLGEK_VMDNTEDH_CGETSQLPDFTTM. RSLLGEK_VMDNTEDH_CGETSQLPDFTTM. RSLLGEK_VMDNTEDH_CGETSQLPDFTTM. RSLLGEK_VMDNTEDH_CGETSQLPDFTTM. RSLLGEK_VMDNTEDH_CGETSQLPDFTTM. RSLLGEK_VMDNTEDH_CGETSQLPDFTTM. RSLLGEK_VMDNTEDH_CGETSQLPDFTTM. RSLLGEK_VMDNTEDH_CGETSQLPDFTTM. RSLLGEK_VMDNTLFNHCETSQLPDFTTM. RSLLGEK_VMDNTLFNHCETSQLPDFTTM. RSLLGEK_VMDNTLFNHCETSQLPDFTTM. RSLLGEK_VMDNTLFNHCETSQLPDFTTM. RSLLGEK_VMDNTLFNHCETSQLPDFTTM. RSLLGEK_VMDNTLFNHCETSQLPDFTTM. RSLLGEK_VMDNTLFNHCETSQLPDFTTM. RSLLGEK_VMDNTLFNHCETSQLPDFTTM. RSLLGEK_VMDNTLFNHCETSQLS_SES_SES_SES_SDS_DSK_SSD_DSK_SSD_DSK_SSS_SSS_SSS_SS	1		ĺ				
GMIWLRHRYPELERPIKVINLALPYFEILACLELIAVSFWATPWASDFITILSCALPYFFCU	1		l			1	
LIANSFWKTPWSVASDFTIILSGLPVYFPGO	ł						
1135	ł	,			ł	1	
1135	ľ						
DLVAFEDVANNFTOEEWSLLDPSQKNLYREV MQETLRNLASIGEKWKDQNIEDQYKYPRNL	1135	2495	Λ	0216	40	410	W WKNKPK WAPPGHLSPRPSCVRSSCMVVPQ
MQETLRNLASIGEKWKDQNIEDQYKNPRNIL RSLLGERVDENTEENHGETSSQIPDDTLNK RISLGERVDENTEENHGETSSQIPDDTLNK RISLGERVDENTEENHGETSSQIPDDTLNK A 9223 3	1133	2403	A	9210	40	410	
RSLLGERYDENTEENHGETSSQIPDDTLINK	1						
1136	İ						
DLVPAKMKGYPHWPARIDDIADGAVKPPPN KYPIFFFCTHETAFLGFKDLFFYDKCKDKYGK KYPIFFFCTHETAFLGFKDLFFYDKCKDKYGK FNRKKGFNEGLWEIQNNPHASYSAPPYSSSD SEAPEANPADGSDADEDDEGRGWAVTAVT ATAASDRWESDSDSDKSNGGLKRKTPALK MSVSKRARKASSDLDQASVSPSEEENSESSSE SEKTSDQDFTPEKKAAVRAPRGPLGGRKKK APSASDSDKAASDGAKPEPVAMARSASSSSS SSSSDSDVVKKPPRGRKPAEKPLPKPRGRK PKPERPPSSSSD KPRLECROPVTVNCTLNLPGSKNAPTTASQK RFERPPSSSSSD KPRLECROPVTVNCTLNLPGSKNAPTTASQK RGSREPPAWA TRSVGRIGHPTNFFVKTGFRCSQAGLKL RGSREPPAWA RGSREP	1126	2406	A	0000		002	
	1130	2400	^	7243	3	983	
PNKKKGFNEGLWEIQNPHASYSAPPYSSSD SEAPEANPADGSDADEDGRGWMAVTAVT ATAASDRMESDSDKSDKSGRGWMAVTAVT ATAASDRMESDSDKSSDNSGLKRKTPALK MSVSKRARKASSDLDQASVSPSEEDSESSSS SEKTSDQDFTPEKKAAVAPRRGPLGGKKK APSASDSDKADSDGAKPEPVAMARSASSSSS SSSSSDSVVKKPPRGKAEKPLPKPRGRK PKPERPSSSSD PROMETICAL P							
SEAPEANPADGSDADEDDEGRYMAVTAVT							
ATAASDRMESDSDSDKSSDNSGLRRKTPALK MSVSKRARKASSDLDQASVSPSEENSESSE SEKTSDQDFTPEKKAAVRAPRRGPLGGRKKK APSASDSDSKADSDGAKPEPVAMARSASSSSS SSSSSDDSVKKKPPRAPKPLPKPRGRK PKPERPPSSSSD FYFRLECRDPVTVNCTLNLPGSKNAPTTASQV GSTWHYRGGLPHPTNFFVKTGFRCSQAGLKL RGSREPPAWA GSTWHYRGGLPHPTNFFVKTGFRCSQAGLKL RGSREPPAWA GSTWHYRGGLPHPTNFFVKTGFRCSQAGLKL RGSREPPAWA GSTWHYRGGLPHPTNFFVKTGFRCSQAGLKL RGSREPPAWA GSTWHYRGGLPHPTNFFVKTGFRCSQAGLKL RGSREPPAWA GSTWHYRGGLPHPTNFFVKTGFRCSQAGLKL RGSREPPAWA GSTWHYRGGLPHPTNFFVKTGFRCSQAGLKL RGSREPPAWA GSTWHTEEGVLVVNVTWRNKTYVGTLLD CTKHDWAPPRFCESPTSDLEMRGGRGRGK ARSAAAAFGSEASTEGQINNRGGARGK ARSAAAAAFGSEASTEGQINNRAGGARGK ARSAAAAAFGSEASTEGQINNRAGGARGK GRRGSLNASGRTFPNCAAEDIKASPSSTNKR KNKPPMELDLNSSSEDNKFPGKRVRTNRSGTP TTPQGKPETITLDQGCSSPVLIDCPHPNCNKK YKHINGLRYHQAHAHLDPENKLEFEPDSEDK ISDCEGGLSNVALECSPSTSVSAYDQLKAPA SPGAGNPPGTPKGKRELMSNGPGSIIGAKAGK NSGKKKGLNNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEABGLDIKKNLGPKEKKKK KNGKGKKGLNNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEABGLDIKKNLGPKEKKKK KNGKGKKGLNNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEABGLDIKKNLGPKEKKKK KLKDKEGKETGSPKMDAKLGKLEDSKGASK NSGKKKGLNNELNNLPVISNMTAALDSCSAA DFTATFITTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	}			1			
MSVSKRARKASSDLDQASVSPSEENSESSSE SEKTSDQDFTPEKKAAVAPARRGPLGGRKKK	į						
SEKTSDQDFTPEKKAAVRAPRGPLGGRKKK AP\$ASDSDSKADSDGAKPEPVAMARSASSSSS SSSSDSDVSVKKPPRGKRPAEKPLPKPRGRK PKPERPFSSSSSD PFPELECRDPVTVNCTLNLPGSKNAPTTASQV GSTWNYRGGLPHPTNFFVKTGFRCSQAGLKL RGSREPPAWA TRSVGVNTCEVGVVTEPECLGPCEPGTSVNL EGIVWHETEEGVLVVNVTWRNKTYVGTLLD CTKHDWAPPRFCESPTSUSAGRRGRGRGKR AR\$AAAAPGSEASFTESRGLQNKNRGGANGK GRRGSLNASGRRTPPNCAAGDIKASPSSTNKR KNRPPMELDLNSSSEDNKEGKNYAFTNGSATP TTPQGKPETTFLDQGCSSPVLIDCPHPNCNKK YKHINGLRYHQAHAHLDPENKLEFEPDSEDK ISDCEEGLSNVALECSPSTSVSAYDQLKAPA SPGAGNPPGTPKGKRELMSNGPGSIIGAKAGK ANCKTDKNPSKLESARPITAPAPAPPPPQLIA PFTATFTTTTGTIGFLSLTTTVVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKKERR KLKDKEGKEGKETGSPKMDAKLGKLEDSKGASK DLPGHFLKDHLNKNEGLANGLSSQESRMAS IKARADKVYTFTDNAPSSIGS TRRGGPWRRTAAAGILPGKAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMOVSSLVIGALL HEQEMOVSSLVIGALL 1140 2490 A 9238 248 328 MAQGNNYGGTSNGVADESPNMLVYRKV 1141 2491 A 9242 2 535 FVEAAVKMLGSLVIRKRALAPRLLIRLIRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGRFDTKCLAA ATWGRLPGFEETLPQQDSWNGVPSRAGIGM WPWAAALVYHCYSKSPSNKDAALLEARAQ NMQEVSRNRCALLHSAAVQEYGYGN 1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLRMAQGCD FALGNDFLNITTKAGATGRGGGRFDTKCLAA ATWGRLPGFEETLPQQDSWNGVPSRAGIGM WPWAAALVYHCYSKSPSNKDAALLEARAQ NMQEVSRNRCALLHSAAVQEYGYGN HLCFWFFVGLFLPEQQIMLFATLRMAQGCD FALGNDFLNITTKAGATGRGGGRFDTKCLAA ATWGRLPGFEETLPQQDSWNGVPSRAGIGM WPWAAALVYHCYSKSPSNKDAALLEARAQ NMQEVSRNRCALLHSAAVQEYGYGN HLCFWFFVGLFLPEQQIMLFATLRMAQGCD FALGNDFLNITTKAGATGRGGGRFDTKCLAA ATWGRLPGFEETLPQQDSWNGVPSRAGIGM WPWAAALVYHCYSKSPSNKDAALLEARAQ NMQEVSRNRCALLHSAAVQEYGYGN HLCFWFFVGLFLPEQQIMLFATLRMAQGCD FALGNDFLNITTKAGATGRGGGRFDTKCLAA ATWGRLPGFEETLPQQDSWNGVPSRAGIGM TLCFWFFVGLFLPEQQIMLFATLLRMAQGCD FALGNDFLNITTCAGT GICKEYSRQ GICKEYSRQ GLYGTSTIQPPGAPDG GLYGYSTIQPPGAPDG GLYGYSTIQPPGAPDG GLYGYSTIQPPGAPDG GLYGYSTIQPPGAPDG GLYGYSTIQPPGAPDG GLYGYSTIQPPGAPDG GLYGYSTIQPPGAPDG GLYGYSTIQPPGAPDG GLYGYSTIQPPGAPDG GLYGYSTIQPPGAPDG GLYGY	i						
APSASDSDSKADSDGAKPEPVAMARSASSSSS SSSSSDVSVKKPPRGRKPAEKPLPKPRGRK SSSSSDVSVKKPPRGRKPAEKPLPKPRGRK FKPERPPSSSSSD SSSSSDVSVKKPPRGRKPAEKPLPKPRGRK FKPERPPSSSSSD SSSSSDVSVKKPPRGRKPAEKPLPKPRGRK FKPERPPSSSSSD SSSSDVSVKKPPRGRKPAEKPLPKPRGRK FKPERPPSSSSSD SSSSDVSVKKPPRGRCPKPKPKFGRCSQAGLKL GSTWPYRGGLPHPTNFFVKTGFRCSQAGLKL RGSREPPAWA GSTWPYRGGLPHPTNFFVKTGFRCSQAGLKL RGSREPPAWA FTSVGVNTCEVGVVTEPECLGPCEPGTSVNL EGIVWHETEEGULVNNVTWRNKTYVGTLLD CTKHDWAPPRFCESPTSDLEMRGGRGRGKR ARSAAAAPGSEASFTESRGLQNKNRGGANGK GRRGSLNASGRRTPPNCAAEDIKASPSTNKR KNRPPMELDLNSSSEDNKFGKRVATNSGSTPL TTPQGKPETTFLDQGCSSPVLIDCPHPNCNKK YKHINGLRYHQAHALDPENKLEFEPDSEDK ISDCEEGLSNVALECSEPSTSVSAYDQLKAPA SPGAGNPPOTPKGKRELMSNOPGGSIGAKAGK ANNCKTDKNPSKLRSARPIAPPAPIPPQLIA PPTATFTTTTTGTIPGLPSLTTTVVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKKEKR KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHFLKDHLNKNEGLANGLSESQESRMAS KAFADKVYTFTDNAPSPSIGS KLKDKEGKEGKEKGKKAANCKTDKNPSKLASARPIAPPAPIPPQLIA PPTATFTTTTTGTIPGLPSLTTTVVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKKKEKR KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHFLKDHLNKNEGLANGLSESQESRMAS KAFADKVYTFTDNAPSPSIGS TRGQPWRRRAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVJGAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVJGAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVJGAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVJGAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVJGAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVJGAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGAALQIKTLLRILTILRSP TLRGGGGASGRNVTGSLGEPQWIRVATGG RPGTSPALFSGRGAATGGQGGRFDTKCLAA ATWGRLPGFEETLPQQDSWNQVPSRAGLGM WPWAAAALVVHCYSKSPSNKDAALLEAARAQ NMQEVSRAVALLHSAAVQEYGYGN HLCFWFFCGLFLPEQQIMITATLLRMAQGCD FALGNDFLNITTKAQATKEKLDKLDFIKKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ GLOCKEYSRQ GLOCKEYSRQ GLOCKEYSRQ GLOCKEYSRQ GLOCKEYSRQ GLOCKEYSRQ GLOCKEYSRQ GLOCKEYSRQ GLOCKEYSRQ GLOC							
SSSSSDSDVSVKKPPRGRKPAEKPLPKPRGRK PKPERPPSSSSSD						ļ	
PKPERPPSSSSD							
1137	· .						
1138 2488 A 9231 1664 2 TRSVGVNTCEVGVVTEPECLGPCEPGTSVNL	1127	0407		2000			
RGSREPPAWA	1137	2487	A	9229	21	239	
1138	· .						
EGIVWHETEEGVLVVNVTWRNKTYVGTLLD	1120	0400		0001			
CTKHDWAPPRFCESPTSDLEMRGGRGKR ARSAAAPGSEASFTESRGLQNKNRGGANGK GRRGSLNASGRRTPPNCAAEDIKASPSTNKR KNKPPMELDLNSSSEDNKPGKRVRTNSRSTP TTPQGKPETTIFLDQGCSSPVLIDCPHPNCNKK YKHNGLRYHQAHAHLDPENKLEFEPDSEDK ISDCEEGLSNVALECSEPSTSVSAYDQLKAPA SPGAGNPPGTPKGKRELMSNGPGSIIGAKAGK NSGKKKGLNNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKK ANNCKTDKNIPSISMTAALDSCSAA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKK ANNCKTDKNIPSISLKSARPIAPAPAPTPPQLIA IPTATFITTITGTIPGLPSLTTTVVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKEKR KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHIFLKDHLNKNEGLANGLSESQESRMAS KABADKVYTFTDNAPPSIGS TRRGQPWRRRAAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL HEQEMGVSSLVIGALL 1140 2490 A 9238 248 328 MAQGNNYGGTSNGVADESPNMLVYRKV HEQEMGVSSLVIGALL A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLLRLRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGRFDTKCLAA ATTWRLPGPEETLPGQDSWNGVPSRAGLGM WPWAAALVVHCYSKSPSNKDAALLEAARAQ NMQEVSRNRCALLHSAAVQEYGYGN MPWAAALVVHCYSKSPSNKDAALLEAARAQ NMQEVSRNRCALLHSAAVQEYGYGN TLCGHGGASGRNVTTGKACATGRQGGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRG GLPQQTSTIQPPGTPDGARDFTSTIQPPGAPDG GICKEYSRG GLPQQTSTIQPPGTPDGARDFTSTIQPPGAPDG GICKEYSRG CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRG GLPQQTSTIQPPGTPDGARDFTSTIQPPGAPDG GICKEYSRG CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRG GLPQQTSTIQPPGTPDGARDFTSTIQPPGAPDG GICKEYSRG CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRG GLPQQTSTIQPPGTPDGARDFTSTIQPPGAPDG GICKEYSRG GLPQQTSTIQPPGTPDGARDFTSTIQPPGAPDG GICKEYSRG CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRG GLPQQTSTIQPPGTPDGARDFTSTIQPPGAPDG GICKEYSRG CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRG GLPQQTSTIQPPGTPDGARDFTSTIQPPGAPDG GICKEYSRG CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRG GLPQQTSTIQPPGTPDGARDFTSTIQPPGAPDG GICKEYSRG GLPQQTSTIQPPGTPDGARDFTSTIQPPGAPDG GICKEYSRG GLPQQTSTIQPPGTPDGARDFTSTIQPPGAPDG GICKEYSRG GLPQQTSTIQPPGTPDGARDFTSTIQPPGAPDG GICKEYSRG GLPQQTTTIQPGTPD	1138	2488	Α	9231	1664	2	
ARSAAAAPGSEASFTESRGLQNKNRGGANGK GRRGSLNASGRRTPPNCAAEDIKASPSSTNKR KNRPPMELDLNSSSEDNKPGKRVATTNSRSTP TTPQGKPETTFLDQGCSSPVLIDCPHPNCNKK YKHINGLRYHQAHAHLDPENKLEFEPDSEDK ISDCEEGLSNVALECSEPSTSVSAYDQLKAPA SPGAGNPPGTPKGKRELMSNGPGSIIGAKAGK NSGKKKGLNNELNLPVISNMTAALDSCSAA DGSLAAEMPKLEABGLIDKKNLGDKEKGKK ANNCKTDKNPSKLKSARPIAPAPAPTPPQLIA PTATFTTTTGTIPGLPSLTTTVVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKKEKR KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHFLKDHLNKNEGLANGLSESQESRMAS IKAEADKVYTFTDNAPSPSIGS 1139 2489 A 9234 207 443 TRRGQPWRRRAAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLASC HEQEMGVSSLVIGALL 1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV 1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLLRLIRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGFDTKCLAA ATWGRLPGPEETLPGQDSWNGVPSRAGLGM WPWAAALVVHCYSKSPSNKDAALLEAARAQ NNGEVSRNRCALLHSAAVQEYGYGN 1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ 1143 2493 A 9247 264 115 GLPQQTSTIQPPGTPDGARDFTSTIQPPGAPDG							
GRRGSLNASGRRTPPNCAAEDIKASPSSTNKR KNKPPMELDLNSSSEDNKFGKRAVATNSRSTP TTPQGKPETTFLDQGCSSPVLIDCPHPNCNKK YKHINGLRYHQAHAHLDPENKLEFEPDSEDK ISDCEEGLSNVALECSEPSTSVSAYDQLKAPA SPGAGNPPGTPKGKRELMSNGPGSIIGAKAGK NSGKKKGLNNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEAEGLIDKKNI.GDKEKGKK ANNCKTDKNPSKLKSARPIAPAPAPTPPQLIA PPTAFTTITTGTIPGLPSLTTTVVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKKEKR KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHFIKDHLNKNEGLANGLSESQESRMAS IKAEADKVYTFTDNAPSPSIGS IKAEADKVYTFTDNAPSPSIGS UTAAVSGLLVGYELGIISGALQIKTLLALSC HEQEMGVSSLVIGALL 1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV 1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLLRLRSP TLRGHGGASGRNVTTGSLGEPQMLRVATGG RPGTSPALFSGRGAATGGRQGGFDTKCLAA ATWGRLPGPEETLPGQDSWNGVPSRAGLGM WPWAAALVVHCYSKSPSNKDAALLEAARAQ NNGEVSRNCALLHSAAVQEYGYGN 1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLRMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ 1143 2493 A 9247 264 115 GLPQQTSTIQPPGTPDGARDFTSTIQPPGAPDG				1			
TREAD TO THE COLOR OF THE COLOR	i i			1 . 1]		
TTPQGKPETTFLDQGCSSPVLIDCPHPNCNKK YKHINGLRYHQAHAHLDPENKLEFEPDSEDK ISDCEEGLSNVALECSEPSTSVSAYDQLKAPA SPGAGNPPGTPKGKRELMSNGPGSIIGAKAGK NSGKKKGLNNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEAEGLIDKKNI.GDKEKGKK ANNCKTDKNPSKLKSARPIAPAPAPTPPQLIA IPTATFTTTTTGTIPGLPSLTTTVVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKKEKR KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHFLKDHLNKNEGLANGLSESQESRMAS IKAEADKVYTFTDNAPSPSIGS ITRGQPWRRRAAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL 1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV 1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRKALAPRLLLRLLRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGRFDTKCLAA ATWGRLPGPEETLPGQDSWNGVPSRAGLGM WPWAAALVVHCYSKSPSNKDAALLEAARAQ WNMQEVSRNRCALLHSAAVQEYGYGN 1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ 1143 2493 A 9247 264 115 GLPQQTSTIQPPGTPDGARDFTSTIQPPGAPDG				├			
YKHINGLRYHQAHAHLDPENKLEFEPDSEDK ISDCEEGLSNVALECSEPSTSVSAYDQLKAPA SPGAGNPGTPKGKRELMSNGPGSIIGAKAGK NSGKKGLNNELNILPVISNMTAALDSCSAA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKK ANNCKTDKNPSKLKSARPIAPAPAPTPPQLIA IPTATFTITTGTIPGLPSLTITVVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKKEKR KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHFLKDHLNKNEGLANGLSESQESRMAS IKAEADKVYTFTDNAPSPSIGS NEAEADKVYTFTDNAPSPSIGS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL 1140							
ISDCEEGLSNVALECSEPSTSVSAYDQLKAPA SPGAGNPPGTPKGKRELMSNGPGSIIGAKAGK NSGKKKGLNNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKK ANNCKTDKNVPSKLKSARPIAPAPAPTPPQLIA IPTATFTTTTTGTIPGLPSLTTTVVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKKEKR KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHFLKDHLNKNEGLANGLSESQESRMAS IKAEADKVYTFTDNAPSPSIGS 1139 2489 A 9234 207 443 TRRGQPWRRRAAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL 1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV 1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLLRLRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGRFDTKCLAA ATWGRLPGPEETLPGQDSWNGVPSRAGLGM WPWAAALVVHCYSKSPSNKDAALLEAARAQ NMQEVSRNRCALLHSAAVQEYGYGN 1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ 1143 2493 A 9247 264 115 GLPQQTSTIQPPGTPDGARDFTSTIQPPGAPDG	1						
SPGAGNPPGTPKGKRELMSNGPGSIIGAKAGK NSGKKKGLNNELNILPVISNMTAALDSCSAA DGSLAAEMPKLEAEGLIDKKNI.GDKEKGKK ANNCKTDKNPSKLKSARPIAPAPAPPPPQLIA IPTATFTTTTTGTIPGLPSLTTTVVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKEKR KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHFLKDHLNKNEGLANGLSESQESRMAS IKAEADKVYTFTDNAPSPSIGS 1139 2489 A 9234 207 443 TRRGQPWRRRAAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL 1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV 1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLIRLIRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGFDTKCLAA ATWGRLPGPEETLPGQDSWNGVPSRAGLGM WPWAAALVVHCYSKSPSNKDAALLEAARAQ NMQEVSRNRCALLHSAAVQEYGYGN 1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ 1143 2493 A 9247 264 115 GLPQQTSTIQPPGTPDGARDFTSTIQPPGAPDG							YKHINGLRYHQAHAHLDPENKLEFEPDSEDK
NSGKKKGLNNELNILPVISNMTAALDSCSAA DGSLAAEMPKLEAEGLIDKKNI.GDKEKGKK ANNCKTDKN\PSKLKSARPIAPAAPTPPQLIA IPTATFTTTTGTIPGLPSLTTTVVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKKEKR KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHFLKDHLNKNEGLANGLSESQESRMAS IKAEADKYYTFTDNAPSPSIGS TRRGQPWRRAAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL 1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV 1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLIRLILRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGRFDTKCLAA ATWGRLPGPEETLPGQDSWNGVPSRAGLGM WPWAAALVVHCYSKSPSNKDAALLEAARAQ VNMQEVSRNRCALLHSAAVQEYGYGN 1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ 1143 2493 A 9247 264 115 GLPQQTSTIQPPGTPDGARDFTSTIQPPGAPDG	!						
DGSLAAEMPKLEAEGLIDKKNI.GDKEKGKK ANNCKTDKNPSKLKSARPIAPAPAPTPPQLIA IPTATFTITTGTIPGLPSLTTTVVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKKEKR KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHFLKDHLNKNEGLANGI.SESQESRMAS IKAEADKVYTFTDNAPSPSIGS 1139 2489 A 9234 207 443 TRRGQPWRRAAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL 1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV 1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRKALAPRLLIRLIRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGRFDTKCLAA ATWGRLPGPEETLPGQDSWNGVPSRAGLGM WPWAAALVVHCYSKSPSNKDAALLEAARAQ NMQEVSRNRCALLHSAAVQEYGYGN 1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ 1143 2493 A 9247 264 115 GLPQQTSTIQPPGTPDGARDFTSTIQPPGAPDG	[[i 1			
ANNCKTDKNPSKLKSARPIAPAPAPTPPQLIA IPTATFTTTTTGTIPGLPSLTTTVVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKKEKR KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHFLKDHLNKNEGLANGLSESQESRMAS IKAEADKVYTFTDNAPSPSIGS 1139 2489 A 9234 207 443 TRRGQPWRRAAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL 1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV 1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRKALAPRLLIELLRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGRFDTKCLAA ATWGRLPGPEETLPGQDSWNGVPSRAGLGM WPWAAALVVHCYSKSPSNKDAALLEAARAQ \(\text{NMQEVSRNRCALLHSAAVQEYGYGN}\) 1142 2492 A 9245 157 466 HLCFWFFVGFLFLPEQQIMLFATLLRMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ 1143 2493 A 9247 264 115 GLPQQTSTIQPPGTPDGARDFTSTIQPPGAPDG					Ì	į l	
PTATFTITTTGTIPGLPSLTTTVVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKKEKR KLKDKEGKETGSPKMDAKL.GKLEDSKGASK DLPGHFLKDHLNKNEGLANGLSESQESRMAS IKAEADKVYTFTDNAPSPSIGS 1139 2489 A 9234 207 443 TRRGQPWRRAAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL 1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV 1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLLRLLRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGRFDTKCLAA ATWGRLPGPEETLPGQDSWNGVPSRAGLGM\ WPWAAALVVHCYSKSPSNKDAALLEAARAQ \NMQEVSRNRCALLHSAAVQEYGYGN 1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ 1143 2493 A 9247 264 115 GLPQQTSTIQPPGTPDGARDFTSTIQPPGAPDG	1						
KPIQPKPTIMGEPITVNPALVSLKDKKKKEKR KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHFLKDHLNKNEGLANGLSESQESRMAS IKAEADKVYTFTDNAPSPSIGS 1139 2489 A 9234 207 443 TRRGQPWRRAAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL 1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV 1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLLRLLRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGRFDTKCLAA ATWGRLPGPEETLPGQDSWNGVPSRAGLGM\ WPWAAALVVHCYSKSPSNKDAALLEAARAQ \NMQEVSRNRCALLHSAAVQEYGYGN 1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ 1143 2493 A 9247 264 115 GLPQQTSTIQPPGTPDGARDFTSTIQPPGAPDG	j						
Standard Standard	1	İ	ľ	. !	i	1	LPTATFTTTTGTIPGLPSLTTTVVQATPKSPPL
DLPGHFLKDHLNKNEGLANGLSESQESRMAS IKAEADKVYTFTDNAPSPSIGS 1139 2489 A 9234 207 443 TRRGQPWRRAAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL 1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV 1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRKALAPRLLLRLRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGRFDTKCLAA ATWGRLPGPEETLPGQDSWNGVPSRAGLGM WPWAAALVVHCYSKSPSNKDAALLEAARAQ NMQEVSRNRCALLHSAAVQEYGYGN 1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ 1143 2493 A 9247 264 115 GLPQQTSTIQPPGTPDGARDFTSTIQPPGAPDG			1		l		
IKAEADKVYTFTDNAPSPSIGS IIA39 2489 A 9234 207 443 TRRGQPWRRRAAAAGILPGREAAACLPSC/AS VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL 1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV 1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLRLRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGRFDTKCLAA ATWGRLPGPEETLPGQDSWNGVPSRAGLGM WPWAAALVVHCYSKSPSNKDAALLEAARAQ NMQEVSRNRCALLHSAAVQEYGYGN 1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ 1143 2493 A 9247 264 115 GLPQQTSTIQPPGTPDGARDFTSTIQPPGAPDG			ļ			,	
1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV 1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLLRLRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGRFDTKCLAA ATWGRLPGPEETLPGQDSWNGVPSRAGLGM\ WNQEVSRNRCALLHSAAVQEYGYGN 1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ 1143 2493 A 9247 264 115 GLPQQTSTIQPPGTPDGARDFTSTIQPPGAPDG		ļ					DLPGHFLKDHLNKNEGLANGLSESQESRMAS
VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL 1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV 1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLRLLRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGRFDTKCLAA ATWGRLPGPEETLPGQDSWNGVPSRAGLGM\ WPWAAALVVHCYSKSPSNKDAALLEAARAQ \NMQEVSRNRCALLHSAAVQEYGYGN 1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ 1143 2493 A 9247 264 115 GLPQQTSTIQPPGTPDGARDFTSTIQPPGAPDG	1120	2400		007.4		110	IKAEADKVYTFTDNAPSPSIGS
HEQEMGVSSLVIGALL HEQEMGVSSLVIGALL HEQEMGVSSLVIGALL HEQEMGVSSLVIGALL MAQGNNYGQTSNGVADESPNMLVYRKV 1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLLRLRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGRFDTKCLAA ATWGRLPGPEETLPGQDSWNGVPSRAGLGM\ WPWAAALVVHCYSKSPSNKDAALLEAARAQ NMQEVSRNRCALLHSAAVQEYGYGN 1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ 1143 2493 A 9247 264 115 GLPQQTSTIQPPGTPDGARDFTSTIQPPGAPDG	1139	2489	Α	9234	207	443	TRRGQPWRRRAAAAGILPGREAAACLPSC/AS
1140 2490 A 9238 248 328 MAQGNNYGQTSNGVADESPNMLVYRKV 1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLRLLRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGRFDTKCLAA ATWGRLPGPEETLPGQDSWNGVPSRAGLGM\ WPWAAALVVHCYSKSPSNKDAALLEAARAQ \NMQEVSRNRCALLHSAAVQEYGYGN 1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ 1143 2493 A 9247 264 115 GLPQQTSTIQPPGTPDGARDFTSTIQPPGAPDG		l	1	, 1			
1141 2491 A 9242 2 535 FVEAAVKMLGSLVLRRKALAPRLLLRLRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGRFDTKCLAA ATWGRLPGPEETLPGQDSWNGVPSRAGLGM\ WPWAAALVVHCYSKSPSNKDAALLEAARAQ \NMQEVSRNRCALLHSAAVQEYGYGN 1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ 1143 2493 A 9247 264 115 GLPQQTSTIQPPGTPDGARDFTSTIQPPGAPDG	1140	2400					
TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGFFDTKCLAA ATWGRLPGPEETLPGQDSWNGVPSRAGLGM\ WPWAAALVVHCYSKSPSNKDAALLEAARAQ \NMQEVSRNRCALLHSAAVQEYGYGN 1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ 1143 2493 A 9247 264 115 GLPQQTSTIQPPGTPDGARDFTSTIQPPGAPDG							
RPGTSPALFSGRGAATGGRQGGRFDTKCLAA ATWGRLPGPEETLPGQDSWNGVPSRAGLGM\ WPWAAALVVHCYSKSPSNKDAALLEAARAQ \NMQEVSRNRCALLHSAAVQEYGYGN 1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ 1143 2493 A 9247 264 115 GLPQQTSTIQPPGTPDGARDFTSTIQPPGAPDG	1141	2491	Α	9242	2	535	
ATWGRLPGPEETLPGQDSWNGVPSRAGLGM\ WPWAAALVVHCYSKSPSNKDAALLEAARAQ\ \NMQEVSRNRCALLHSAAVQEYGYGN 1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLIRMAQGCD\ FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC\ CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF\ GICKEYSRQ 1143 2493 A 9247 264 115 GLPQQTSTIQPPGTPDGARDFTSTIQPPGAPDG		Í	ſ	ĺ			TLRGHGGASGRNVTTGSLGEPQWLRVATGG
WPWAAALVVHCYSKSPSNKDAALLEAARAQ \NMQEVSRNRCALLHSAAVQEYGYGN 1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ 1143 2493 A 9247 264 115 GLPQQTSTIQPPGTPDGARDFTSTIQPPGAPDG		ļ	ţ				
1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ 1143 2493 A 9247 264 115 GLPQQTSTIQPPGTPDGARDFTSTIQPPGAPDG			i	ļ			
1142 2492 A 9245 157 466 HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ 1143 2493 A 9247 264 115 GLPQQTSTIQPPGTPDGARDFTSTIQPPGAPDG			l		ł		
FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ 1143 2493 A 9247 264 115 GLPQQTSTIQPPGTPDGARDFTSTIQPPGAPDG			l			l	
FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ 1143 2493 A 9247 264 115 GLPQQTSTIQPPGTPDGARDFTSTIQPPGAPDG	1142	2492	A	9245	157	466	
CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ 1143 2493 A 9247 264 115 GLPQQTSTIQPPGAPDG					ļ	1	FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC
1143 2493 A 9247 264 115 GLPQQTSTIQPPGTPDGARDFTSTIQPPGAPDG			ļ		ļ		
113 GELQQISHQITGIFDGAADI ISHQFGAFDG							GICKEYSRQ
	1143	2493	A	9247	264	115	GLPQQTSTIQPPGTPDGARDFTSTIQPPGAPDG

SEQ ID	SEO ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		į.	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		i		amino acid	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
		1]	peptide	sequence	/=possible nucleotide deletion, \=possible
Ì				sequence		nucleotide insertion
1144	2494	A	9260	1	401	KKVPGRLSEMSFSLNFTLPANTTSSPVT\DCGP
}]		_]	SLGLAAGIPLLVATALLVALLFTLIHRRRSSIE
						AMEESDRPCEISEIDDNPKISENPRRSPTHEKN
	İ					TMGAQEAHIYVKTVAGSEEPVHDRYRPTIEM
						ERRR
1145	2495	Α	9264	175	411	METIWIYQFRLIEIGDSTVGKSCLLHRFTQGRF
						PGLRSPACDPTVGVDFFSRLLEIEPGKRIKLLL
1116	2406		0077	500	0.14	WDTAGQERFISIT
1146	2496	Α	9277	592	814	MFTYLEGREGIKSQPKMEPHSVT\RLECSGMI
						SAHCSLNLPGTSDSPASASR/VAGTTGMRHHA WLIFAFLVETGF
1147	2497	A	9279	1255	2	FRRGRRGEEEKEEEEEEEGWVNGMENSHPP
114/	2471	^	7219	1233	4	HHHHQQPPPQPGPSGERRNHHWRSYKLMIDP
1						ALKKGHHKLYRYDGQHFSLAMSSNRPVEIVE
		1				DPRVVGIWTKNKE\LELSVPKFKIDEFYVDQV
			[PPKQVTFAKLNDNIRENFLRDMCKKYGEVEE
			1			VEILYNPKTKKHLGIAKVVFATVRGAKDAVQ
	}	1	1		}	HLHSTSVMGNIIHVELDTKGETRMRFYEL\LV
		ł	1			TGRYTPQTLPVGELDAVSPIVNETLQLSDALK
1						RLKDGGLSAGCGSGSSSVTPNSGGTPFSQDTA
]]]]]		}	YSSCRLDTPNSYG/QGTPLTPRLGTPFSQDSSY
•		l	į.			SSRQPTPSYLFSQDPAVTFKARRHESKFTDAY
						NRRHEHHYVHNSPAVTAVAGATAAFRGSSD LPFGTVGGTGGSSGPPFKAQPQDSATFAHTPP
						PAQATPAPGFR
1148	2498	Α	9302	1026	6	IASIQNADTMPGVGLLVSHFSTLVSRQRCPNY
			""	1020		ADPONLTDVSIFLLLEVSGDPELQPVLAGLFL
		İ				SMCLVT:VLGNLLIILAISPDSHLHTPMYFFFSN
		ĺ				LSLPDV\GFTSTTVPK\MIVDI\QSRSRVISYAG
						CLTQKSLFAIFGGTEE\NMLLSVMAYDRFVAI
						CHPLYHSAIMNPCFCAFLVLLSFFFLSLLDSQL
		}				HSWIVLQFTIIKNVEISNFVCDPSQLLKFACSD
			1			SIINSIFIYFHKDPERQLVLAGLFLSMCLVTVL GNLIIILDVSPDSHLPTPMYFFLSNLSLPDIGFT
						STTVPKMIVDIQSHGRVIFYAGCLTQMSLFAIF
						GGMEERHAPECDGL
1149	2499	A	9303	<u> </u>	699	MASQEKDIFIGWGTIHLFRKPQRSFFGKLLRE
				-		FRLVAADRSMGRYMLFGVINLICTGFLLMWC
						SSTNSIALT\SYTYLTIFDLFSLMTCLISYWVTL
						RKPSPVYSFGFERLEVLAVFASTVLAQLGALF
					:	ILKESAERFLEQPEIHTGRLLVGTFVALCFNLF
						TMLSIRNKPFAYVSEAASTSWLQEHVADLSR
						SLCGIPGLSSIFLPRMNPFVLIDLAGAFALCIT
1150	2500		0202			YMLIEI
1150	2500	A	9308	797	693	DRSTSVTRAGVQWCSLGSLQPRTPGLLRSSCL
1151	2501	_	9309	205	106	SLP
1171	2301	A	לטנע	203	406	VAIKELPVLWKWSKPTR\TAKEPPQTQQRAG SKTAAPPCOWSRMASEGPNIPCPGARHSDKQ
					·	FLICTI
1152	2502	A	9314	913	504	KPSPLITPPAVVLPPSAVLNLVNTFSSFPQVEV
1		••	/5.7	,	207	QGPLCGPRKGRLAVTIPFFGLS/LPKYMDHRR
			1			PPPHR\EIFFVFLAETGFHRASQAGPDLPTS/S/I
						PPTSA/FPKCWEYRSEPQCLPGCLSFSGILLDL
						GTNVSLRAA
1153	2503	Α	9315	392	1	HPHRPRPGFRSPARSSRPCPVLTSLLPPFPSPSP
						PADDLVKAGRDRKDPQVR/ERRLRPNPGRLG
			·			GPR\PRPARARS/CHQPRLTRVCPRSPPPEARA
						PAPAAPARGRGAPKRNRPRTDTRAPRGSSAR
			L			PGNS

SEQ ID NO: of nucl- eotide seq-	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496	Predicted beginning nucleotide location correspondi	Predicted end nucleotide location corresponding to last amino	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first amino acid residue of peptide sequence	acid residue of peptide sequence	Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
1154	2504	A	9321	331	433	MPCI/QAQYGTPAPSPGPRDHSASDPLTPEFIK PT
1155	2505	Α	9324	180	275	MEEPQSDPSVEPPLSQETFSDLWKLLSENNVL
1156	2506	Α	9326	383	619	MISPSRTEGDPLPLPPÆGEGGEVRGFGGGPAK EAAQRHCRASVSILRMRRPGQGSSRPARVPL RGPDSHRLREPPPSPP
1157	2507	A	9327	152	292	YERRGRSQGGGSHPAGAQPGGRAIGAGWQS KEPLWEGLQRSGSPLPG
1158	2508	A	9328	1	430	QELKQGPNPLAPSPSAPSTSAGLGDCNHRVD LSKTFSVSSALAMLQERRCLYVVLTDSRCFL VCMCFLTFIQALMVSGYLSSVITTIERRYSLKS SESGLLVSCFDIGNLVVVVFVSYFRGRRRRP/ RVAAVGGLLDLEGGEMI
1159	2509	A	9334	108	383	KGNQVNGNGNQLKRKHESMCPVSLTQNTVR LMEAGLPQKQAERADELFEAGLVIYVKLDER VLNAL\YSSVGLQWFKESDLSHLRLLEISFR
1160	2510	A	9338	2	430	FVGRPRGLSDRLEDLFLAGFRVGERLRTAAM KRYVRILLLGEGAEHVADPVPGGRGVPRGEA DHTDQELREEIHKANVERVVHDVSQEATIEKI RTKWIPLV/RWGDHA/EGPVGIKSYLPSGRSM EAELPIMSQLTEIETCVEC
1161	2511	Α	9341	1	390	NSRVDDFVAPGLSEAGKLLGLEFPERQRLAA AVG/CSPMSGVISMSAPFFLGKIIDAIYTNPTV DYSDNLTRLCLGLSGVFLCGAAANAIRVYLM QTSRQRVVKRLRTSLFSSILGQEVAFSDKAGT GELI
1162	2512	A	9343	84	837	QGRFRAFCWQRDFLQPPGMRLSALLALASKV TLPPHYRYGMSPPGSVADKRKNPPWIRRRPV VVEPISDEDWYLFCGDTVEILEGKDAGKQGK VVQVIRQRNWVVVGGLNTHYRYIGKTMDYR
						GTMIPSEAPLLHRQVKLVDPMDRKPTEIEWR FTEAGERVRVSTRSGRIIPKPEFPRADGIVPET WIDGPKDTSVEDALERTYVPCLKTLQEEVME AMGIKETRINTRRSIGIEPGAEQLLPNFCPSLE G
1163	2513	A	9346	967	616	DSLALSPRLECSGAISAHCNLTPPGFTPFSCLS LPSSWAYRCASPHPDNFFVFLVESGFHHVGQ AGLKLLISSDPPTSA/FPKCWDYRRD\SSAPAT FSSYQRNNPDLILNDTIMPNIK
	2514	A	9347	3	1099	SSFPTCMRTVFHSNTSVSSLLHRPGHVTPQLTI HGGWRHHRDHTAIDEWDFNPSKFLIYTCLLL FSVLLPLRLDGIIQWSYWAVFAPIWLWKLLV VAGASVGAGVWARNPRYRTEGEACVEFKA MLIAVGIHLLLLMFEVLVCDRVERGTHFWLL VFMPLFFVSPVSVAACVWGFRHDRSLELEILC SVNILQFIFIALKLDRIIHWPWLVVFVPLWILM SFLCLVVLYYIVWSLLFLRSLDVVAEQRRTH VTMAISWITIVPLLTFEVLLVHRLDGHNTFS YVSIFVPLWLSLLTLMATTFRKGGNHWWF AIRRDF/CQDQLPQPTGKPPPPPLTDHHGEKA LPLQNKDRGSWPASRGSPRLL
1165	2515	A	9362	547	991	DVSIGPPLLRRPCSGREQTRSLSFPSDPESSFSP VPEGVRLADGPGHCKGRVEVKHQNQWYTV CQTGWSLRAAKVVCRQLRCGRAVLT\QKRC TKHAYGRKPIWLSQMACSGPEPTLHDCPFRP LGEDTLFHVEYTSVHGRERLSAKD
1166	2516	A	9363	201	387	PPILRWTPPSGKNFFFFFFESEFY/SSPRVECS GAISAHLAHCNLCLPGSSDSPASAFQVAS
1167	2517	Α	9368	707	1087	AVLTPCLSPCSPSRIPRP\SRPYPGRRSLSHTPP

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion PRPLILYAPAP\RPAGTAFIPHSHPPPPDLLRPT ATPA/TPCPSLPPPPRPLHPTQPSTALLPDPPPW PLPFPPPSS/RPPRPDCSTSYSPTFPPPT
1168	2518	A	9375	511	15	MMLSEETSAVRPQKQTRFNGAKLVWMLKGS PITVTSAVIIVLMLLMM/IFSPWLATHDPNAID LTARLLPPSAAHWFGTDEVGRDLFSRVLVGS QQSILAGLVVVATTGMIGSPLECLFGELGGRA DAIFMRVMDIMRS/IPSLVLTMEKTAALGPSL FNAMQASSEH
1169	2519	A	9377	42	410	GNGRVAPRDPGAVASAEPGLTTHDSGVNPN NSARRMEAMASGSNWLSGVNVVLVMAYWS LVFVLLFIFAKRQIMRFAMKSLRGPHGPVGH NAPKDLKEEIDILLSRVHNIKYEP\HLLADDDA
1170	2520	A	9378	302	1303	GVSGFSASVLRQRRMEDELEPSLRPRTQIQGR ILLLTICAAGIGGTFQFGYNLSIINAPTLHIQEF TNETWQARTGEPLPDHLVLLMWSLIVSLYPL GGLFGALLAGPLAITLGRKKSLL\VNNIFVVS AAILFGFSRKAGSFEMIMLGRLASWGVNAGV SMNIQP\MLPGGESAPKELRGAVAMSSAIFTA LGIVMGQVVGLSTTAATGLRGL\AGELEELEE ERAACQGCRARRPWELFQHRALRRQVTSLV VLGSAMELCGNDSVYAYASSVFRKAGVPEA KIQYAIIGTGSCELLTAVVSVSLEGALPPPAL WGGTPRSFALNQFTLQKKKK
1171	2521	A	9381	2	412	RGPASAQEDERARTAPLERVRARGRMTTSSA LFPSLLPCSWSTSNKYLAEFRAGKMSLKGTTE TPDKRKGLAY/IQQTDDSLIHFCWKDRTSGNV EDDLIIFPDDCEFKRLPQCPNGRVYVLKFKAG SKRLFFWMQEP
1172	2522	A	9384	20	355	GWNGRSTEASPAAEAPHVPHKET\KAAMGTQ CTHGGKVRPPHDMLTIVVHKIKLFVLCHSL LQLCAIMISDYLKSSIYIVEKRLGLFRPTSGLL ASFNEVGNTALIVLESY
1173	2523	Α	9393	430	87	LCQCIVPGQQKETFSLNPSSATVRFYL*LSLQ QRKEDQ*IIL*YHLNKDCLHIFMSAITLYMKI* KIFVLFDFNIMFETPFYII*FIFLFSQNLKRIRQV IRPPISFSKINNGP
1174	2524	A	9397	77	374	ERLEIGRLGGERGSGPASCLRVIDVSGMWDQ RLVKLALLQLLRAFYGIKVKGVRVHRDCGTF ESSSTLIRVS*FGVPCNALAHFGVTHF*YILDF LGML
1175	2525	A	9399		397	HESSRADRDKMDTRGSTYTDADPVNKSGGT AKMNKWSKGKVRDKLNNLVLFDTATYDKL CKEVPNYKLITLAVVSERLKIPGSLARAALHE LLSRGLI*LVIQHIAQVIY
1176	2526	A	9408	2	299	LDLTHVLSLSISLTVTILGTTFGMVIPLLDVVY GERGYAQNGDF*DAQLDDYSFSCYSHAQVN GAPNSLTRAYDDP*VKISGLECQKVGALVEV KCLNL
1177	2527	A	9416	2	402	CNFLRSSRIRVHSTPAASTMPPKVDPNEIKVV YLRCTGGEVRATSALAPKIGPLGLSSIKVGVD FV*ATGDWNVLIISVILTIRILLSHIFVVPPFFCF DHLIAFWDLQSLIFLHVIFSLFITLLLFCFFSIF
1178	2528	A	9419	142	426	TPLFDLWPRVVLSWLETVLTSLRTRRAASGPP ACRIMPTTVDDVLEHGGEVHFLQKQMLYLL ALI*DTFAPIYVGIVFLGFTPDHRCRSPGVAEL
1179	2529	A	9420	1450	1655	LSSAGTKMNLN*KNYWPGASAHACNPSTLG GQSRCITRSGDRDHPG*HGETPSVLKIQKISRA WWRAP

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
1180	2530	A	9422	176	375	HRPQTTRPDWKPRT*PQGK*GRLSSEISPASPP SRFSRSTKPVPPKADPPARQKLTGVLHAPLLK L
1181	2531	Α	9436	2	274	PIAASLRMYNLQPYTEENLICTAFATMVETVP IARTILDRLTGIPHGYCFVE*ADWATADKCVH IYNGKPLPGATPLLSLQLHQLAHLGS
1182	2532	Ā	9442	3	240 :	VDKCSSKSIVLSEYCPHCMCSLSTDPKPFGQL SMILK*MGAGDEKISAMGKARVDHRELYLGL LYPTEDYKLTFRARH
1183	2533	A	9444	384	3	LKDFQPWALHDWPLFCCCTFLLFLVLECFTR KGCSGWAPWLSLQCQHFGRPRWADHLRSGV RDQPGQYSKTTFLPKIQKLAGHSGAHL*S*LL ERMRWKNRLNPGGRSCSEPRWHHCTPGWAT ERG
1184	2534	A	9462	391	655	LSGFKSLMPKIPLQYIYVRVRTTWSFCLPLDG RKLMLS*YSK*LT*KYNILPEYSRMTLPPGMV IHTCNPSTLGGRAGWIV*AQEFET
1185	2535	A	9467	215	566	RCPMWQGQASRMDPAKAKDREASTCCSLA WWWGWECWVRALKLSSGPAGPLACWVAK KKSLSLSGPVYPSEKGAGLYVF*DRVSLCHPG WSAVVQFWLTAASNSCFSLLSSWDYRCA
1186	2536	A	9468	275	452	HIPQLHTKTHYVPTRMVNKI*QIDNSKPWQR GG*TGILTHCW*ESKLVQPLWKIVWHYQ
1187	2537	A	9469	388	3 :	EVAPGPSQILPRRVTDGGDRPQFSLPGPRLPQ SSRGAEPCLSNCIHSPAPRKQRMGDSDQ*STP NPASPHPEAPQEPWDSASGSVGSFSLGRGAK ASS*VPGKGRGPRQGSELLAETILELFLALAN S
1188	2538	Α	9471	124	397	TMDKKNRHGNSLDMASEIHMTGPMCLIENTT GRLMANPEALKILSAITQPMVEEAIAGLYRAC
1189	2539	A	9480	584	769	*FYLTNNLAGMKKGLCLGSTEQAHTIGI GHVQSQHFGRPRRADHLRSGDRDHPG*HDET PSLLKIQKISWAWWRAPVVPATWEAEAEEW R
1190	2540	A	9483	463	86	VTVGLTLLLRGAPRFTAG*PPSGGGPPLAPLL PRQHCTLQTHRHLHPEAPVKV*KT*RLFPGLR GASSCRRRCNPVLAARKAGSPRSHSTRENC RRSRCPDTAHRRRRRGRRRNPSCVRSPRWR
1191	2541	Α .	9489	1	411 ;	LADALCLSAAATGAVRPGARAQPSTRRRLSP SVRVCCRAAAASNLLYSSCLQRHSERASEEG ERGSLSAKCCSLVLRGGCSSSNSHSFRRIT*EI MAAFVLLSYEQRPLKRPRLGPPDVYPPDPKQ KEEELTAVNVK
1192	2542	A	9497	389	161	VSFLSMSSGHCIRSTRGSKMVSWSVIAKIQEI* CEEDERKMAREFLAEFMSTYVMMNIHMIVE KDTYSDHEEINTS
1193	2543	A	9509	186	1	IAKSQ*KRWQRSGAMETLKHGWWECKLVQF FGKTFVNVN*S*TYVYPCDKIILLLGLYPTEM
1194	2544	Α	9512	58	433	PLQRSKCLTLRCLRAKPWAWSQSPRACSSAL LKSSRSRASSLNVQCILQSNPQGHQRI*KQKA SSKGQQFRR*KEHPFMLKTLNKLRIEGT*LKI RRAIYDNPTANIIVEGQKLEAFPLRTGTRO
1195	2545	A	9515	595	1223	GHGAPSFQTQVPRTP*ASWPVVPAASESAPAP AGGGASLPVAAGSCAAAPHTEPGAPQHLLDC PCPLCLARPPRRPLPDTCYGPGSGRSASLAEPP LPRCSCAPLRSASAPQVS*CV*AVNLLPHNL* PLHLLLHD*EKAWGFLFSSASHCFQGQICLLP APGSGPCGATARPSRGGRAGGSRARRPIPPGP GTRRTPSGCQNPAASGG

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	,			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1				peptide		/=possible nucleotide deletion, \=possible
1106	0646	· · · · · · · · · · · · · · · · · · ·	0510	sequence	120	nucleotide insertion
1196	2546	A	9518	229	468	RSPTATPAPHAMGPGAPFARGGRPLPLLGAM
1					į	AERVAPGWDLHTPYLPRTNSRRTPHL**EPHA GYIGALFPMSGGWPGGQ
1197	2547		9521	289	448	IAWLSGLFFPSNQANLCFLCYKLTADSRYRG
1177	2347	Λ.	9321	209	440	HAMRHLTGNTSMAIRFL*ADSRFQVQRARYE
						APNWKYKYGY*IPVDMLC
1198	2548	Ā	9524	204	1	KNKKTTKCLSIVTLNISGPNQ*NKRHRVAEWI
1170	2540	, · ·	7524	204	*	VKQEPNICHL*ETHFPFRDTYRLKEREQKKRK
1					1	SSYS
1199	2549	A	9546	1785	1943	GGRFKESKLTNAGWQRNSFFIGPPKSIPWAA
1					""	V*QRGDGKNPGVTHLNRPVGTX
1200	2550	A	9548	186	1	VNAEKEF*KIQHYFMTKSQNKLHIEHTYLKPI
						KAIYDKWTSDIMLNLQKL*AFFLRVIVRQI
1201	2551	A	9549	591	2	SSVVEFPRGPRSSLPPLDSTFPCGSSPNWTGGC
						GSCPSGE*LVSPGSEQRKKYSNSNVIMHETSQ
						YHVQHLATFIMDKSEAITSVDDAIRKLVQLSS
1				i	İ	KEKIWTQEMLLQVNDQSLRLLDIESQEELEDF
						PLPTVQRSQTVLNQLRYPSVLLLVCQDSEQSK
						PDVHFFHCDEVEAELVHEYMESALTDCRLGK
						AMRP
1202	2552	Α	9552	428	1	KYGNEGHWSRQCPNPGKPIRPCPLCRGPHWK
1						LDCERPPQGPLPSLPELAKTSYSDLTGLATED
1						*WGPGMDAPATTIASSKTRVTLMVAGRPVFF
						LI*YRATYSALPNFSGPTQSSQVSVVGIDGQV
1203	2553	Α	9568	517	738	SKPRATPPLFCSLHTF
1203	2333	А	9508	517	/38	RRKFERKQKQ*RYREGKQYRQRDKMKEWG EKEKRRREKGEREERKMRHRERKGESGQRD
						TMENWRVERLTEKER
1204	2554	A	9573	83	415	EDKRLRLVDGDSRCAGRV*IYHDGFWGTICD
1201	2057	11	7575	. 03	415	DGWDLSDAHVVCQKLGCGVAFNATVSAHFG
1						EGSGPIWLDDLNCTGTESHLWQCPSRGWGQ
1						HDCRHKEDAGVICSEFTALR
1205	2555	. A	9577	64	424	ARGSCPTRPRTANGRMGETKDAPOMLVTFK
						DVAVTFFREEWRQLVLVHRTLYR*GMLETC
						GLLDTLRHNVPQPDVVHLLYHGTQLLIVKRE
						VSHSPCAGDMRELFTREATLTPHPYNNGA
1206	2556	Α	9584	38	476	TLGAVLFSEVSKESSTSHSGGQLGRQNRHPKL
				İ		SNFITPSSPRLKP*TASSQRNLGQILNMFLTAV
						NPQPLSTPSWQIETKYSTKVLTGNWMEERRK
						GLPYKHLITHHQEPPHRYLISTYDDHYNRHG
1207	2557	_	0505		410	YNPGLPPLRTWNGQKLLWL
1207	2557	Α	9586	2	412	LRSSPAALLRALCITTVTGTALALRSRVATTN
						PDGCRNVLRPKYYRLCDKAESWGIALETVPT GVAVTSWAIMLTVLTLVCKGQDYNRRQKLP
		'				THILCLL*EKGIFGLTFAFIIGLDGSTGPTRFFL
						FGILFSICFS
1208	2558	Ā	9597	122	3	IKNYWPGMVAHACNPSPLGGRGRWIA*AQK
	2000	**	,,,,			FADAWADAW
1209	2559	A	9611	148 .	558	KSLRNVWDLLNNTWKADRFFCHSSRTSTIRK
/						GDPGPTFSKMSIWTSGRTSSSYRHDEKRNIYO
						RIRDHDLLDKRKTVTALKAGEDRAILLGLAM
						MVCSIMM*FLLGITLLRSYMQSVWTRESQCT
		1		ļ		LLNASITETFNC
1210	2560	Α	9618	384	2	SLHDMLMLAEQQQKQKWAVNTQNTAWSNA
						DSKFGQRILEKMEWSKGRGLGVQEQGGPDDI
						KVQVKNNDLGLQATINNEANWIAHQDDFNW
		[[LLAELNTCQRQETADS***WSPKNSHVGKDS
						GELSAK
1211	2561	<u>A</u>	9620	316	610	QKHPGGGQLGRSPQEDSRFHNKASSGVSRVR

No. of No. of No. of nucleotide continue co	OFO ID	I OF O TO	17.4	T 686	T N	T. S. 11	
	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Decision Sequence USSN Decation 1944 Sequence 1944 Sequence 1944 Sequence 1944 Sequence 1945			noa	1			
Sequence							
Uniform Section Sect					1		
mmino acid residue of peptide requence mmino acid residue of peptide requence mmino acid residue of peptide requence mmino acid requence mmino a		uence	ĺ	1			
Page Page	uonoo		l	/14			
	,					sequence	/=nossible nucleotide deletion \=nossible
LGRAWWLTPVPTLWEAKAGGSPEPD*AGRE			l			J	
							
1212 2562 A 9623 297 344 GPPUGDTYQKERITQI-FQAGNIS.CIAMTR TREL*KGGGKGRHP-AVVPELKKGGYGVKA. ALI.NTSNCT-CP-ETKMLSDDPKACVPEVSSA DL*NTSFQVIR 1213 2563 A 9624 2 356 AELSLASTAGGRNIS.GDELPDYDRAPISSPILA TSGTILSAISCL WDL.PTPVLRVGLSCQPSNSSS IPRMYSTDVEAAVNSLEDLY.LQAYYAYLCVC LYPHRDMALECVSRPL*BLS 1214 2564 A 9634 776 912 SLSRWVRAKI.*VPYNGENCINPRGGGSEPR SHYCTPAWATEKDS SHYCTPAWATEKDS 1215 2565 A 9636 220 426 KPGNFAVSSEY*DITSGGLKTAVRG*IBMTST 1216 2566 A 9637 391 76 GFRIGREN HOLDGENGFLDKT*KAQATKAKI DK GENERATIR FRANKINSPDDTIKENVTISNIRTRKI NHLPETERNLLEHGLMYIRLNAAFCSLVAHS LFGHLKAT LFCHIK					ļ	1	GSRL*SOHFGRPRRVDHLRSAVODOPGOHGE
1212 2562 A 9623 297 344 GPPUDGDVQKIEKTIOLEQAQNLSLCLAMTR TREL*KGGGGKIE*AVPFIKKGGYGVKAA ALMITSNCT*CF*ETKMLSDDPKACVFEVSSA DL*NTSFQVIR ALMITSNCT*CF*ETKMLSDDPKACVFEVSSA DL*NTSFQVIR ALMITSSQUAN ALMITSNCT*CF*ETKMLSDDPKACVFEVSSA DL*NTSFQVIR AELSLASTAGGRNTSGDSLPDYDRAPISSPLA TSGTILSAISCLWDLPTPVLRVGLSCQPSMSSC IPPRMYSTDVBAAVNSLED!VQAYVAYLCVC LYFHRDDMALEGVSRFI-*ELAE 1214 2564 A 9634 776 912 SLSRWVRAKLVPFVNQENCLNPRGGGCSEPR SHYCTPAWATEKDS SHYCTPAWATEKDS			1	ļ	}]	
TRELF-KGGGGRAIEF-A-VVPFLKKGGYGVKAA ALLNTSNCT*CF*ETKMLSDDPKACVFEVSSA ALLNTSNCT*CF*ETKMLSDDPKACVFEVSSA ALLNTSNCT*CF*ETKMLSDDPKACVFEVSSA ALLSLASTAGGRNTSGDSLPDYDRAPISSPLA TSGTILSAISCL WDLPTPVLRVGLSCQPSMSSS PRMYSTDVEAAVNSLEDLYLQAYYAYLCVC LYFHRDDMALBGVSRFL*BY LYFHRDDMALBGVSRFL*BY LYFHRDDMALBGVSRFL*BY LYFHRDDMALBGVSRFL*BY LYFHRDDMALBGVSRFL*BY LYFHRDDMALBGVSRFL*BY LYFHRDDMALBGVSRFL*BY LYFHRDDMALBGVSRFL*BY LYFHRDDMALBGVSRFL*BY LYFHRDDMALBGVSRFL*BY LYFHRDDMALBGVSRFL*BY LYFHRDDMALBGVSRFL*BY LYFHRDDMALBGVSRFL*BY LYFHRDDMALBGVSRFL*BY LYFHRDDMALBGVSRFL*BY LYFHRDDMALBGVSRFL*BY LYFHRDDMALBGVSRFL*BY LYFHRDDMALBGVSRFL*BY LYFHRDDMALBGYSRFL*BY LYFHRDDMALBGYSRFL*BY LYFHRDMALBGYSRFL*	1212	2562	A	9623	297	344	
ALINTSNCT*CF*FIKMLSDDPKACVFEVSSA DL*NTSFGVIR							
1213 2563 A 9624 2 356			1		}	l	
TSGTILSAISCLWDLPTPVLRVGLSCQPSMSSC IPRMYSTDVBAAVNSLEDLYLQAYYAYLCV IPRMYSTDVBAAVNSLEDLYLQAYYAYLCV LYFHRDDMALGGVSRF1-ELAE							
1214 2564 A 9634 776 912 SLSRWVRAKLYVPYNQENCLYQAYYAYLCVC LYFHRDDMALEGVSRFL*ELAE 1215 2565 A 9636 220 426 SLSRWVRAKL*VPYNQENCLYPRGGGCSEPR SHYCTPAWATEKDS 1216 2566 A 9637 391 76 CFLEDGCTQAS*AEEAAVSFSMAEEQGTSTS 1216 2566 A 9637 391 76 CFLEDGCTQAS*AEEAAVSFSMAEEQGTSTS 1217 2567 A 9655 2008 2432 LHCKMGALETQHTPCSQNMLRSLQKCCCKV 1218 2568 A 9658 3 405 HASARALSPNLSPNIKMAISGGPVLGFFIIA 1218 2568 A 9658 3 405 HASARALSPNLSPNIKMAISGGPVLGFFIIA 1219 2569 A 9662 3 284 PHWTEKRYMQLGSRYSTYLVPLSFLYWRLE*LARLD 1219 2569 A 9662 3 284 PHWTEKRKMQDTGSILPIHWFGFGYAALVA 1210 VIMASAQEPWAIKEHVIIQAGEFYLINPOSGEF 1210 2570 A 9669 200 699 LLLTGYIQTLQNQLSGNQQEMQAVDNLTSA 1220 2571 A 9669 200 699 LLLTGYIQTLQNQLSGNQQEMQAVDNLTSA 1221 2571 A 9676 164 562 KERDSTFSAAMTTMQGMEQAMPGAGPGVP 1222 2572 A 9688 43 412 VAKMYKCCSAIGCASRCLPNSKLKGLTFHVF 1223 2573 A 9696 308 564 RTSMGILYSEPICAAYVQDGRIPLICACKGRHVIIVSFL 1224 2574 A 9700 3 632 DAWASGELGSGLGSLGSGRFPINSKLKGLTFANVIIVSTL 1224 2574 A 9700 3 632 DAWASGELGSGLGSGSTPNSKLKGLTFHVSITALL	1213	2563	A	9624	2	356	
IPRMYSTDVEAAVNSLEDLYLQAYYAYLCVC				! :			
LYPHRDDMALEGVSRFL** LAE 1214 2564 A 9634 776 912 SLSRWPKAKL** VPYNQENCLNPRGGGSEPR 1215 2565 A 9636 220 426 KPONFAVSSEY* DITSGQLKTAVGG* EMTST 1216 2566 A 9637 391 76 CFLEGGTQAS**AEEAAVSPSMAEEEQGSTSC 1217 2567 A 9655 2008 2432 LHCKMGALETQTHPCSQNMLRSLQKCCKV 1218 2568 A 9655 2008 2432 LHCKMGALETQTHPCSQNMLRSLQKCCKV 1218 2568 A 9658 3 405 HASARALLSPNLSPNNKMAISGGPVLGFFILA 1219 2569 A 9662 3 284 PDWTEKRMQDTGSILPLHWEGFGYAALVA 1219 2569 A 9662 3 284 PDWTEKRMQDTGSILPLHWEGFGYAALVA 1210 2570 A 9669 200 699 LLLTGYIGTLQNQCLSGNQEMQAVDNLTSA 1220 2570 A 9669 200 699 LLLTGYIGTLQNQCLSGNQEMQAVDNLTSA 1221 2571 A 9676 164 562 KERDSSTFSAAMTTMQGMEQAMPGAGPGVP 1222 2572 A 9688 43 412 VAKMVKCCSAIGCASRCLPNSKLKGLTFIVF 1223 2573 A 9696 308 564 RTSMGILYSEPPOLACKING NOR PRINCIPLE 1224 2574 A 9700 3 632 DAWASGGLGSLEPROLACKING NOR PRINCIPLE 1224 2574 A 9700 3 632 DAWASGGLGSLEPNOTRACKTRE 1224 2574 A 9700 3 632 DAWASGGLGSLEPNOTRACKTRE 1224 2574 A 9700 3 632 DAWASGGLGSLEPNOTRACKTRE 1226 1226 1227 22576 A 9700 3 632 DAWASGGLGSLEPNOTRACKTRE 1227 2574 A 9700 3 632 DAWASGGLGSLEPNOTRACKTRE 1228 1229 12						Į.	IPRMYSTDVEAAVNSLEDLYLOAYYAYLCVG
1214 2564 A 9634 776 912 SLSRWYRAKLI-VPYNQENCLNPRGGGCSEPR SHYCTPAWATEKDS 1215 2565 A 9636 220 426 KPONFAVSSEY*DITSGQLKTAVRG*IEMTST EENFGEKLHDIGFGNGFLDKT*KAQATKAKI DK 1216 2566 A 9637 391 76 CFLEDGCTQAS*AEEAAVSPSMAEEEQGSTSG RERRSIRFKMKNHSPDDTIKENVTISNIRTRKI NHLPETERNILEHELGMYIRLNAAFCSLVAHS LFOFILKAT LHCKMGALETQTHPCSQNMLRSLQKCCCKV EHHLQPVQVLQTLLHSATAGTGCRRPARPP PAPPTPTBWRSRQSGKQSERAS*LKGRGRYGI GALGGRGGRALGGSGMAPPELGGCHLFSGCKH RRRRGSDAAPGEEAGT HASARALLSPNLSPNNKMAISGGPVLGFFIIA VLMSAQEPWAIKEEHVIIQAEFYLNPDQSGEF MLDFEGEDTHFGDMAKKETVWRLE*LARLD NFEAQRALANIAADQAALEIMDMGSDYTLIP NVPPKVTVL 1219 2569 A 9662 3 284 PPWTEKRKMQDTGSILPLHWFGFGYAALVA YGGIIGVYKAGSVPSLAAGLLFGSLSGLGAYQ LSQDPRNVWVFLATSGTLAGIMGMRFYHSG KL KL STOPP STANDAR ST			ł			· ·	LYFHRDDMALEGVSRFL*ELAE
SHYCTPAWATEKDS	1214	2564	A	9634	776	912	SLSRWVRAKL*VPYNQENCLNPRGGGCSEPR
1216							
1216	1215	2565	Α	9636	220	426	KPGNFAVSSEY*DITSGQLKTAVRG*IEMTST
1216						'	EENFGEKLHDIGFGNGFLDKT*KAQATKAKI
RERRSIRFKMKNHSPDDTIKENYTISNIRTRKI			L				DK
1217 2567 A 9655 2008 2432	1216	2566	A	9637	391	76	CFLEDGCTQAS*AEEAAVSPSMAEEEQGSTSC
1217 2567 A 9655 2008 2432						i	
1217	1		1	İ			NHLPETERNLLEHGLMYIRLNAAFCSLVAHS
EEHHLQPVQVLQTLLHSATAGTGCRRPARPP PAPPTPTPWRSRQSGKQSERAS*LKGRGRYGI GALGGRGGRALGGSRWPPPLPGETLFSGCKH RRRRRGSDAAPGEEAGT RRRRRGSDAAPGEEAGT WLMSAGEPWAIKEHVIIQAEFYLNPDQSGEF MLDFFGEDTFHGDMAKKETVWRLE*LARLD NFEAQRALANIAADQAALEIMDMGSDYTLIP NVPPKVTVL							1 -T
PAPPTPTPWRSRQSGKQSERAS*LKGRGRYGI	1217	2567	A	9655	2008	2432	
1218 2568			ĺ	İ	İ	1	
1218			!				PAPPTPTPWRSRQSGKQSERAS*LKGRGRYGL
1218							
VLMSAQEPWAIKEEHVIIQAEFYLNPDQSGEF MLDFEGEDTFHGDMAKKETVWRLE*LARLD NFFAQRALANIAADQAALEIMDMGSDYTLIP NVPKVTVL		0.000	L.,.				
MLDFEGEDTFHGDMAKKETVWRLE*LARLD NFEAQRALANIAADQAALEIMDMGSDYTLIP NFEAQRALANIAADQAALEIMDMGSDYTLIP NVPPKVTVL	1218	2568	Α	9658	3	405	
NFEAQRALANIAADQAALEIMDMGSDYTLIP NVPPKVTVL 1219 2569 A 9662 3 284 PDWTEKRKMQDTGSII.PLHWFGFGYAALVA YGGIIGYVKAGSVPSLAAGLLFGSLSGLGAYQ LSQDPRNVWVFLATSGTLAGIMGMRFYHSG KL 1220 2570 A 9669 200 699 LLLTGYIQTLQNQQLSGNQQEMQAVDNLTSA PGNTSLCTRDYKITQVLFPLLYTVLFFVGLITN GLAMRIFFQIRSKSNFIIFLKNTVISDLLMILTF PFKILSDAKLGTGPLRTFVCQVTSVIFYFTMYI SISFLGLITIDRYQKTTRPFKTSNPKNLLGAKIL K 1221 2571 A 9676 164 562 KERDSSTFSAAMTTMQGMEQAMPGAGPGVP QLGNMAVIHSHLWKGLQEKFLKGEPKVLGV VQLTALMSLSMGITMMCMASNTYGSNPISV YIGYTIWGSVMFIISGSLSIAAGIRTTKGLVRG SLGMNITSS 1222 2572 A 9688 43 412 VAKMVKCCSAIGCASRCLPNSKLKGLTFHVF PTDENIKRKWVLAMKRLDVNAAGIWEPKKG DVLCSRHFKKTDFDRSAPNIKLKPGVIPSIFDS PYHLQGKREKLHCRKNFTLKTVPATNYNH 1223 2573 A 9696 308 564 RTSMGILYSEPICQAAYQNDFGQVWRWVKE DSSYANVQDGFNGDTPLICACRRGHVRIVSFL LKKECLCQPQKPERENLLALCCE 1224 2574 A 9700 3 632 DAWASGGELGSLFDHHVQRAVCDTRAKYRE							
1219 2569 A 9662 3 284 PPWTEKRKMQDTGSILPLHWFGFGYAALVA	i		ľ	ĺ			
1220 2570 A 9669 200 699 LLLTGYIQTLQNQQLSGNQQEMQAVDNLTSA PGNTSLCTRDYKITQVLFPLLYTVLFFVGLITN GLAMRIFFQIRSKSNFIIFLKNTVISDLLMILTF PFKILSDAKLGTGPLRTFVCQVTSVIFYFTMYI SISPLGLITIDRYQKTTRPFKTSNPKNLLGAKIL K 1221 2571 A 9676 164 562 KERDSSTFSAAMTTMQGMEQAMPGAGPGVP QLGNMAVIHSHLWKGLQEKFLKGEPKVLGV VQILTALMSLSMGITMMCMASNTYGSNPISV YIGYTIWGSVMFIISGSLSIAAGIRTTKGLVRG SLGMNITSS 1222 2572 A 9688 43 412 VAKMVKCCSAIGCASRCLPNSKLKGLTFHVF PTDENIKRKWVLAMKRLDVNAAGIWEPKKG DVLCSRHFKKTDFDRSAPNIKLKPGVIPSIFDS PYHLQGKREKLHCRKNFTLKTVPATNYNH 1223 2573 A 9696 308 564 RTSMGILYSEPICQAAYQNDFGQVWRWVKE DSSYANVQDGFNGDTPLICACRRGHVRIVSFL LKKECLCQPQKPERENLLALCCE 1224 2574 A 9700 3 632 DAWASGGELGSLFDHHVQRAVCDTRAKYRE]		ļ				
YGGIIGYVKAGSVPSLAAGLLFGSLSGLGAYQ LSQDPRNVWVFLATSGTLAGIMGMRFYHSG KL 1220 2570 A 9669 200 699 LLLTGYIQTLQNQQLSGNQQEMQAVDNLTSA PGNTSLCTRDYKITQVLFPLLYTVLFFVGLITN GLAMRIFFQIRSKSNFIIFLKNTVISDLLMILTF PFKILSDAKLGTGPLRTFVCQVTSVIFYFTMYI SISFLGLITIDRYQKTTRPFKTSNPKNLLGAKIL K 1221 2571 A 9676 164 562 KERDSSTFSAAMTTMQGMEQAMPGAGPGVP QLGNMAVIHSHLWKGLQEKFLKGEPKVLGV VQILTALMSLSMGITMMCMASNTYGSNPISV YIGYTIWGSVMFIISGSLSIAAGIRTTKGLVRG SLGMNITSS 1222 2572 A 9688 43 412 VAKMVKCCSAIGCASRCLPNSKLKGLTFHVF PTDENIKRK WVLAMKRLDVNAAGIWEPKKG DVLCSRHFKKTDFDRSAPNIKLKPGVIPSIFDS PYHLQGKREKLHCRKNFTLKTVPATNYNH 1223 2573 A 9696 308 564 RTSMGILYSEPICQAAYQNDFGQVWRWVKE DSSYANVQDGFNGDTPLICACRRGHVRIVSFL LKKECLCQPQKPERENLLALCCE 1224 2574 A 9700 3 632 DAWASGGELGSLFDHHVQRAVCDTRAKYRE	1210	2560	Α	0667		-004	
LSQDPRNVWVFLATSGTLAGIMGMRFYHSG KL	1217	2309	^	3002]	204	VGGIGVVKAGEVDELAAGI LEGELEGI GAVO
SECTION PROPERTY			ļ				I SODDDNIVWVEI ATSCTI ACIACADEVISC
1220 2570 A 9669 200 699 LLLTGYIQTLQNQQLSGNQQEMQAVDNLTSA PGNTSLCTRDYKITQVLFPLLYTVLFFVGLITN GLAMRIFFQIRSKSNFIIFLKNTVISDLLMILTF PFKILSDAKLGTGPLRTFVCQVTSVIFYFTMYI SISFLGLITIDRYQKTTRPFKTSNPKNLLGAKIL K 1221 2571 A 9676 164 562 KERDSSTFSAAMTTMQGMEQAMPGAGPGVP QLGNMAVIHSHLWKGLQEKFLKGEPKVLGV VQILTALMSLSMGITMMCMASNTYGSNPISV YIGYTIWGSVMFIISGSLSIAAGIRTTKGLVRG SLGMNITSS 1222 2572 A 9688 43 412 VAKMVKCCSAIGCASRCLPNSKLKGLTFHVF PTDENIKRK WVLAMKRLDVNAAGIWEPKKG DVLCSRHFKKTDFDRSAPNIKLKPGVIPSIFDS PYHLQGKREKLHCRKNFTLKTVPATNYNH 1223 2573 A 9696 308 564 RTSMGILYSEPICQAAYQNDFGQVWRWVKE DSSYANVQDGFNGDTPLICACRGHVRIVSFL LKKECLCQPQKPERENLLALCCE 1224 2574 A 9700 3 632 DAWASGGELGSLFDHHVQRAVCDTRAKYRE							
PGNTSLCTRDYKITQVLFPLLYTVLFFVGLITN GLAMRIFFQIRSKSNFIIFLKNTVISDLLMILTF PFKILSDAKLGTGPLRTFVCQVTSVIFYFTMYI SISFLGLITIDRYQKTTRPFKTSNPKNLLGAKIL K 1221 2571 A 9676 164 562 KERDSSTFSAAMTTMQGMEQAMPGAGPGVP QLGNMAVIHSHLWKGLQEKFLKGEPKVLGV VQILTALMSLSMGITMMCMASNTYGSNPISV YIGYTIWGSVMFIISGSLSIAAGIRTTKGLVRG SLGMNITSS 1222 2572 A 9688 43 412 VAKMVKCCSAIGCASRCLPNSKLKGLTFHVF PTDENIKRK WVLAMKRLDVNAAGIWEPKKG DVLCSRHFKKTDFDRSAPNIKLKPGVIPSIFDS PYHLQGKREKLHCRKNFTLKTVPATNYNH 1223 2573 A 9696 308 564 RTSMGILYSEPICQAAYQNDFGQVWRWVKE DSSYANVQDGFNGDTPLICACRGHVRIVSFL LKKECLCQPQKPERENLLALCCE 1224 2574 A 9700 3 632 DAWASGGELGSLFDHHVQRAVCDTRAKYRE	1220	2570	A	9669	200	600	1
GLAMRIFFQIRSKSNFIIFLKNTVISDLLMILTF PFKILSDAKLGTGPLRTFVCQVTSVIFYPTMYI SISFLGLITIDRYQKTTRPFKTSNPKNLLGAKIL K 1221 2571 A 9676 164 562 KERDSSTFSAAMTTMQGMEQAMPGAGPGVP QLGNMAVIHSHLWKGLQEKFLKGEPKVLGV VQILTALMSLSMGITMMCMASNTYGSNPISV YIGYTIWGSVMFIISGSLSIAAGIRTTKGLVRG SLGMNITSS 1222 2572 A 9688 43 412 VAKMVKCCSAIGCASRCLPNSKLKGLTFHVF PTDENIKRK WVLAMKRLDVNAAGIWEPKKG DVLCSRHFKKTDFDRSAPNIKLKPGVIPSIFDS PYHLQGKREKLHCRKNFTLKTVPATNYNH 1223 2573 A 9696 308 564 RTSMGILYSEPICQAAYQNDFGQVWRWVKE DSSYANVQDGFNGDTPLICACRGHVRIVSFL LKKECLCQPQKPERENLLALCCE 1224 2574 A 9700 3 632 DAWASGGELGSLFDHHVQRAVCDTRAKYRE			••	1007	200	000	
PFKILSDAKLGTGPLRTFVCQVTSVIFYFTMYI SISFLGLITIDRYQKTTRPFKTSNPKNLLGAKIL K 1221 2571 A 9676 164 562 KERDSSTFSAAMTTMQGMEQAMPGAGPGVP QLGNMAVIHSHLWKGLQEKFLKGEPKVLGV VQILTALMSLSMGITMMCMASNTYGSNPISV YIGYTIWGSVMFIISGSLSIAAGIRTTKGLVRG SLGMNITSS 1222 2572 A 9688 43 412 VAKMVKCCSAIGCASRCLPNSKLKGLTFHVF PTDENIKRKWVLAMKRLDVNAAGIWEPKKG DVLCSRHFKKTDFDRSAPNIKLKPGVIPSIFDS PYHLQGKREKLHCRKNFTLKTVPATNYNH 1223 2573 A 9696 308 564 RTSMGILYSEPICQAAYQNDFGQVWRWVKE DSSYANVQDGFNGDTPLICACRGGHVRIVSFL LKKECLCQPQKPERENLLALCCE 1224 2574 A 9700 3 632 DAWASGGELGSLFDHHVQRAVCDTRAKYRE							
SISFLGLITIDRYQKTTRPFKTSNPKNLLGAKIL K 1221 2571 A 9676 164 562 KERDSSTFSAAMTTMQGMEQAMPGAGPGVP QLGNMAVIHSHLWKGLQEKFLKGEPKVLGV VQILTALMSLSMGITMMCMASNTYGSNPISV YIGYTIWGSVMFIISGSLSIAAGIRTTKGLVRG SLGMNITSS 1222 2572 A 9688 43 412 VAKMVKCCSAIGCASRCLPNSKLKGLTFHVF PTDENIKRKWVLAMKRLDVNAAGIWEPKKG DVLCSRHFKKTDFDRSAPNIKLKPGVIPSIFDS PYHLQGKREKLHCRKNFTLKTVPATNYNH 1223 2573 A 9696 308 564 RTSMGILYSEPICQAAYQNDFGQVWRWVKE DSSYANVQDGFNGDTPLICACRRGHVRIVSFL LKKECLCQPQKPERENLLALCCE 1224 2574 A 9700 3 632 DAWASGGELGSLFDHHVQRAVCDTRAKYRE	İ						
1221 2571 A 9676 164 562 KERDSSTFSAAMTTMQGMEQAMPGAGPGVP QLGNMAVIHSHLWKGLQEKFLKGEPKVLGV VQILTALMSLSMGITMMCMASNTYGSNPISV YIGYTIWGSVMFIISGSLSIAAGIRTTKGLVRG SLGMNITSS 1222 2572 A 9688 43 412 VAKMVKCSAIGCASRCLPNSKLKGLTFHVF PTDENIKRKWVLAMKRLDVNAAGIWEPKKG DVLCSRHFKKTDFDRSAPNIKLKPGVIPSIFDS PYHLQGKREKLHCRKNFTLKTVPATNYNH 1223 2573 A 9696 308 564 RTSMGILYSEPICQAAYQNDFGQVWRWVKE DSSYANVQDGFNGDTPLICACRRGHVRIVSFL LKKECLCQPQKPERENLLALCCE 1224 2574 A 9700 3 632 DAWASGGELGSLFDHHVQRAVCDTRAKYRE			l	[
1221 2571 A 9676 164 562 KERDSSTFSAAMTTMQGMEQAMPGAGPGVP QLGNMAVIHSHLWKGLQEKFLKGEPKVLGV VQILTALMSLSMGITMMCMASNTYGSNPISV YIGYTIWGSVMFIISGSLSIAAGIRTTKGLVRG SLGMNITSS 1222 2572 A 9688 43 412 VAKMVKCCSAIGCASRCLPNSKLKGLTFHVF PTDENIKRKWVLAMKRLDVNAAGIWEPKKG DVLCSRHFKKTDFDRSAPNIKLKPGVIPSIFDS PYHLQGKREKLHCRKNFTLKTVPATNYNH 1223 2573 A 9696 308 564 RTSMGILYSEPICQAAYQNDFGQVWRWVKE DSSYANVQDGFNGDTPLICACRGGHVRIVSFL LKKECLCQPQKPERENLLALCCE 1224 2574 A 9700 3 632 DAWASGGELGSLFDHHVQRAVCDTRAKYRE	l						
QLGNMAVIHSHLWKGLQEKFLKGEPKVLGV VQILTALMSLSMGITMMCMASNTYGSNPISV YIGYTIWGSVMFIISGSLSIAAGIRTTKGLVRG SLGMNITSS 1222 2572 A 9688 43 412 VAKMVKCCSAIGCASRCLPNSKLKGLTFHVF PTDENIKRK WVLAMKRLDVNAAGIWEPKKG DVLCSRHFKKTDFDRSAPNIKLKPGVIPSIFDS PYHLQGKREKLHCRKNFTLKTVPATNYNH 1223 2573 A 9696 308 564 RTSMGILYSEPICQAAYQNDFGQVWRWVKE DSSYANVQDGFNGDTPLICACRGHVRIVSFL LKKECLCQPQKPERENLLALCCE 1224 2574 A 9700 3 632 DAWASGGELGSLFDHHVQRAVCDTRAKYRE	1221	2571	Α	9676	164	562	
VQILTALMSLSMGITMMCMASNTYGSNPISV YIGYTIWGSVMFIISGSLSIAAGIRTTKGLVRG SLGMNITSS 1222 2572 A 9688 43 412 VAKMVKCCSAIGCASRCLPNSKLKGLTFHVF PTDENIKRKWVLAMKRLDVNAAGIWEPKKG DVLCSRHFKKTDFDRSAPNIKLKPGVIPSIFDS PYHLQGKREKLHCRKNFTLKTVPATNYNH RTSMGILYSEPICQAAYQNDFGQVWRWVKE DSSYANVQDGFNGDTPLICACRGGHVRIVSFL LKKECLCQPQKPERENLLALCCE 1224 2574 A 9700 3 632 DAWASGGELGSLFDHHVQRAVCDTRAKYRE	i						
YIGYTIWGSVMFIISGSLSIAAGIRTTKGLVRG SLGMNITSS							VQILTALMSLSMGITMMCMASNTYGSNPISV
SLGMNITSS 1222 2572 A 9688 43 412 VAKMVKCCSAIGCASRCLPNSKLKGLTFHVF PTDENIKRK WVLAMKRLDVNAAGIWEPKKG DVLCSRHFKKTDFDRSAPNIKLKPGVIPSIFDS PYHLQGKREKLHCRKNFTLKTVPATNYNH 1223 2573 A 9696 308 564 RTSMGILYSEPICQAAYQNDFGQVWRWVKE DSSYANVQDGFNGDTPLICACRGHVRIVSFL LKKECLCQPQKPERENLLALCCE 1224 2574 A 9700 3 632 DAWASGGELGSLFDHHVQRAVCDTRAKYRE	ł		1				
1222 2572 A 9688 43 412 VAKMVKCCSAIGCASRCLPNSKLKGLTFHVF PTDENIKRK WVLAMKRLDVNAAGIWEPKKG DVLCSRHFKKTDFDRSAPNIKLKPGVIPSIFDS PYHLQGKREKLHCRKNFTLKTVPATNYNH 1223 2573 A 9696 308 564 RTSMGILYSEPICQAAYQNDFGQVWRWVKE DSSYANVQDGFNGDTPLICACRRGHVRIVSFL LKKECLCQPQKPERENLLALCCE 1224 2574 A 9700 3 632 DAWASGGELGSLFDHHVQRAVCDTRAKYRE			•				1
DVLCSRHFKKTDFDRSAPNIKLKPGVIPSIFDS PYHLQGKREKLHCRKNFTLKTVPATNYNH 1223 2573 A 9696 308 564 RTSMGILYSEPICQAAYQNDFGQVWRWVKE DSSYANVQDGFNGDTPLICACRRGHVRIVSFL LKKECLCQPQKPERENLLALCCE 1224 2574 A 9700 3 632 DAWASGGELGSLFDHHVQRAVCDTRAKYRE	1222	2572	Α	9688	43	412	VAKMVKCCSAIGCASRCLPNSKLKGLTFHVF
DVLCSRHFKKTDFDRSAPNIKLKPGVIPSIFDS PYHLQGKREKLHCRKNFTLKTVPATNYNH 1223 2573 A 9696 308 564 RTSMGILYSEPICQAAYQNDFGQVWRWVKE DSSYANVQDGFNGDTPLICACRRGHVRIVSFL LKKECLCQPQKPERENLLALCCE 1224 2574 A 9700 3 632 DAWASGGELGSLFDHHVQRAVCDTRAKYRE	}						
PYHLQGKREKLHCRKNFTLKTVPATNYNH 1223 2573 A 9696 308 564 RTSMGILYSEPICQAAYQNDFGQVWRWVKE DSSYANVQDGFNGDTPLICACRRGHVRIVSFL LKKECLCQPQKPERENLLALCCE 1224 2574 A 9700 3 632 DAWASGGELGSLFDHHVQRAVCDTRAKYRE							
DSSYANVQDGFNGDTPLICACRRGHVRIVSFL LKKECLCQPQKPERENLLALCCE 1224 2574 A 9700 3 632 DAWASGGELGSLFDHHVQRAVCDTRAKYRE	•			[PYHLQGKREKLHCRKNFTLKTVPATNYNH
DSSYANVQDGFNGDTPLICACRRGHVRIVSFL LKKECLCQPQKPERENLLALCCE 1224 2574 A 9700 3 632 DAWASGGELGSLFDHHVQRAVCDTRAKYRE	1223	2573	Α	9696	308	564	RTSMGILYSEPICQAAYQNDFGQVWRWVKE
1224 2574 A 9700 3 632 DAWASGGELGSLFDHHVORAVCDTRAKYRE							DSSYANVQDGFNGDTPLICACRRGHVRIVSFL
1224 2574 A 9700 3 632 DAWASGGELGSLFDHHVORAVCDTRAKYRE					İ	i	LKKECLCQPQKPERENLLALCCE
langua comunication de la comuni	1224	2574	A	9700	3	632	DAWASGGELGSLFDHHVQRAVCDTRAKYRE
							GRRPRAVKVYTINLESQYLLIQGVPAVGVMK
ELVERFALYGAIEQYNALDEYPAEDFTEVYLI	I					ł	ELVERFALYGAIEQYNALDEYPAEDFTEVYLI
KFMNLQSARTAKRKMDEQSFFGGLLHVCYA							KFMNLQSARTAKRKMDEQSFFGGLLHVCYA
PEFETVEETRKKLQMRKAÝVVKTTENKDHY							
VTKKKLVTEHKDTEDFRQDFHSEMSGFCKA		l				Ì	
ALNTSAGNSNPYLPYSCELPLCYFSSK	I	1					ALNTSAGNSNPYLPYSCELPLCYFSSK

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
1225	2575	A	9710	sequence	163 .	nucleotide insertion RSGCVLRMTEWETGAPAVAETPDIKLFGKWS TDDVHINDISLQDYIAGVRLILL
1226	2576	A	9713	82	492	QGLPSFLPAFGPSGSWLGPAPTLGSSCNTVDT ICHGYSEIRPLFYLSFCDLLLGLCWLTETLLYG ASVANKDIICYNLQAVGQIFYISSFLYTVNYI WYLYTELRMKHTQSGQSTSPLVIDYTCRVCQ MAFVFSSLI
1227	2577	A	9720	3	416	GKWKRTQVPLLGEECADMDLARKEFLRGNG LAAGKMNISIDLDTNYAELVLNVGRVTLGEN NRKKMKDCQLRKQQNENVSRAVCALLNSGG GVIKAEVENKGYSYKKDGIGLDLENSFSNML PFVPNFLDFMQNGNYF
1228	2578	A	9723	278	411	EASSSNTVASNVADKTDPHSMNSRVFIGNLN TLVLQKSDVEAVF
1229	2579	Α	9725	121	902	LFAMSGFENLNTDFYQTSYSIDDQSQQSYDY GGSGGPYSKQYAGYDYSQQGRFVPPDMMQP QQPYTGQIYQPTQAYTPASPQPFYGNNFEDEP PLLEELGINFDHIWQKTLTVLHPLKVADGSIM NETDLAGPMVFCLAFGATLLLAGKIQFGYVY GISAIGCLGMFCLLNLMSMTGVSFGCVASVL GYCLLPMILLSSFAVIFSLQGMVGIILTAGIIG WCSFSASKIFISALAMEGQQLLVAYPCALLYG VFALISVF
1230	2580	Α	9739	11	247	TFVLNMNTPKEEFQDWPIVRIAAHLPDLIVYG HFSPERPFMDYFDGVLMFVDISGKCKRDVCL MWMSNRLAWEFTCRA
1231	2581	A	9744	37	1100	TPLFDFWPGFVLSWLQPLSASLRARRAASGPP ACRIMPTTVDDVLEHGGEFHFFQKQMFFLLA LLSATFAPIYVGIVFLGFTPDHRCRSPGVAELS LRCGWSPAEELNYTVPGPGPAGEASPRQCRR YEVDWNQSTFDCVDPLASLDTNRSRLPLGPC RDGWVYETPGSSIVTEFNLVCANSWMLDLFQ SSVNVGFFIGSMSIGYIADRFGRKLCLLTTVLI NAAAGVLMAISPTYTWMLIFRLIQGLVSKAG WLIGYILITEFVGRRYRRTVGIFYQVAYTVGL LVLAGVAYALPHWRWLQFTVALPNFFFLLY YWCIPESPRWLISQNKNAEAMRIIKHIAKKNG KSLPASL
1232	2582	A	9753	164	517	PGPGMQGPPPITPTSWSLPPWRAYVAAAVLC YINLLNYMNWFIIAGVLLDIQEVFQISDNHAG LLQTVFVSCLLLSAPVFGYLGDRHSRKATMS FGILLWSGAGLSSSFISPRYSWLF
1233	2583	Α	9757	25	419	LPAPWTERVRKSEGLVGTCLGDPMASPRTVT IVALSVALGLFFVFMGTIKLTPRLSKDAYSEM KRAYKSYVRALPLLKKMGINSILLRKSIGALE VACGIVMTLVPGRPKDVANFFLLLLVLAVLF FHQLV
1234	2584	A	9765	71	456	RLELDWGFSLHFLPVAYLCPLSSGFEMNVQP CSRCGYGVYPAEKISCIDQIWHKACFHCEVC KMMLSVNNFVSHQKKPYCHAHNPKNNTFTS VYHTPLNLNVRTFPEAISGIHDQEDGEQCKSV FHWD
1235	2585	Α	9767	52	559	IRSGAMSVDKAELCGSLLTWLQTFHVPSPCA SPQDLSSGLAVAYVLNQIDPSWFNEAWLQGI SEDPGPNWKLKVTSGLLIRGQTGEEMTRDGP ARHMSWVMGRKRDRCLVINHLFIHSSMEYSP CARPGHSARNNTDKNLPHTAIILVTSNTYTTI KINFQAGRSGSCL
1236	2586	Α	9770	352	608	FRGEALTVRFLTKRFIGEYASNFESIYKKHLC

NO of peptide much color of peptide side ide of peptide side side of peptide side side of peptide side side of peptide side side of peptide side side side of peptide side side side of peptide side side side of peptide side side side side side of peptide side side side of peptide side side side side side side side s	SEQ ID	SEQ ID	Met	SEQ	Predicted	Daniel de de ma	
			1	1		Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
			""	1 .			
1237 2587 A 9793 266 515 Controlled 1238 2588 A 9802 537 967 ELGAGROSE E			1	1			I-Icolousing K-Lusing I - Lousing
1237 2587 A 9793 266 515 MILAUTEPRICE			1				M=Methionine N=Assemaine D=Pooline
minino acid residue of peptide residue of peptide residue of peptide sequence T-Intronine, W-Valine, W-ITyptophan, Y-ITyrosine, X-Unknown, Y-Sup codon, /-possible nucleotide deletion, \possibl							O=Glutamine P=Argining S=Socies
Peptide of peptide sequence				-• ·			TeThrennine Veveline WeTerstenhan
Peptide							V=Tyrosine Y=Inknown *=Stan codon
1237 2587 A 9793 266 515 DERENDENSSTATAKALI 1237 2587 A 9793 266 515 DIVENTISPENSTOKKEFSLTSELHWA 1238 2588 A 9802 537 967 ELGAGRSDREAMELAVKEEISVEDEAVDENI 1238 2588 A 9802 537 967 ELGAGRSDREAMELAVKEEISVEDEAVDENI 1240 2590 A 9805 105 540 VEDPAMVRAGAVGAHLPASGLDIFGDLKK 1240 2590 A 9819 3 305 DIGENPLEAMELAVGAMPRIREDILLIPHE 1241 2591 A 9834 841 1269 SPARGKSRETHARSGLOPHELK 1242 2592 A 9843 3 589 TISCOPALPHENSTSTATORAKMENSLAAMENSCALLIPHEN 1244 2594 A 9846 198 411 WISHIRAGKMITUK 1245 2595 A 9848 116 650 PIGGER PLAY RANGAGRARER CHAPTER CHAP		İ	1	1		Sequence	/=nossible nucleotide deletion \-nossible
	İ	ĺ	ì			i	
1237 2587 A 9793 266 515 NILLINIPPERJELIZISSISSISSISSIANALI			 		sequence		
1237					i		DCEALANDELDLC2ALAWAR2F12EFHMA
1238 2588 A 9802 537 967 ELGAGRSIDEAMEAAVKEISVEDEAVDKNI FRDCNKIAFYRRQKOWLSKKSTYRALLDSVT FRDCNKIAFYRRQKOWLSKKSTYRALLDSVT FRDCNKIAFYRRQKOWLSKKSTYRALLDSVT FRDCNKIAFYRRQKOWLSKKSTYRALLDSVT FRDCNKIAFYRRQKOWLSKKSTYRALLDSVT FRDCNKIAFYRRQKOWLSKKSTYRALLDSVT FRDCNKIAFYRRQKOWLSKKSTYRALLDSVT FRDCNKIAFYRRQKOWLSKKSTYRALLDSVT FRDCNKIAFYRRQKOWLSKKSTYRALLDSVT FRDCNKIAFYRRQKOWLSKKSTYRALLDSVT FRDCNKIAFYRRQKOWLSKKSTYRALLDSVT FRDCNKIAFYRRQKOWLSKKSTYRALLDSVT FRDCNKIAFYRRQKOWLSKKSTYRALLDSVT FRDCNKIAFYRRQKOWLSKKSTYRALLDSVT FRDCNKIAFYRRQKOWLSKKSTYRALLDSVT FRDCNKIAFYRQKOWLSKSTYRALLDSVT FRDCNKIAFYRQKOWLSKSTYRALLDSVT FRDCNKIAFYRQKOWLSKSTYRALLDSVT FRDCNKIAFYRQKOWLSKSTYRALLDSKT FRDCNKIAFYRQKYRQKOWLSKSTYRALLDSKT FRDCNKIAFYRQKYRQKYYLSGSMEPAFRIRGDLEFT.TYRGKOWLSKYRKYT FRDCNKIAFYRQKYYLSGSMEPAFRIRGDLEFT.TYRGKYRKYRQKYYLSGSMEPAFRIRGDLEFT.TYRGKYRKYRQKYYLSGSMEPAFRIRGDLEFT.TYRGKYRKYRQKYYLSGSMEPAFRIRGDLEFT.TYRGKYRKYRQKYYLSGSMEPAFRIRGDLEFT.TYRGKYRKSAF FORDEN PRAGRETYR GESPIVVYNGGSPYK FORDEN PRAGRETYR GESPIVYNJERSKYRSSAF FROKKINGT FROM PRAGRETYR GESPIVYNGGST FROKKINGT FROM PRAGRETYR GESPIVYNJERSKYRSSAF FROKKINGT FROM PRAGRETYR GESPIV FROKKINGT FROM PRAGRETYR GESPIVYNJERSKYR GESPIV FROKKINGT FROM PRAGRETYR GESPIV FROKKINGT FROM PRAGRETYR GESPIV FROKKINGT FROM PRAGRETYR GESPIV FROKKINGT FROM PRAGRETYR GESPIV FROM PRAGRETYR GESPI							
1238	1	===,			200] 313	
1238	Į.				1	!	
1239 2589 A 9805 105 540 VPCDPAMPKAGAVGAHLPASGLDIFGDLKK MNRRQLYYQVLNFAMIVSSALMIWKGILVL GESPIT GESPIT GES	1238	2588	A	9802	1 527	067	
1239 2589 A 9805 105 540 VPODPAMYRAGAVGAHLPASGLDIFGDLIKK MPKBQLYQVLNFAMIVSSALMIWKGLIVLT GSESPIVVVLSGSMEP AFRICOLLFI, THE PIRAGEIVYFKVEGRDIPIYHKIVFRED PIRAGEIVYFKVEGRDIPIYHKIVFRED PIRAGEIVYFKVEGRDIPIYHKIVFRED PIRAGEIVYFKVEGRDIPIYHKIVFRED PIRAGEIVYFKVEGRDIPIYHKIVFRED PIRAGEIVYFKVEGRDIPIYHKIVFRED PIRAGEIVYFKVEGRDIPIYHKIVFRED PIRAGEIVYFKVEGRDIPIYHKIVFRED PIRAGEIVYFKVEGRDIPIYHKIVFRED PIRAGEIVFKYEVGROPPIYHRVIKVHEKDNG DIKFLTKGDNNEGDDRGSYK 1240 2590 A 9819 3 305 TDCRPDFLPCAARRGGGGCCGAGWVAEWS PQPLDPAMILWMQGFVLEAVACQDNDDYLR YGLEFDLDCNGGOVVDIELGGLRNWSSAF DPNSEEHG SPARGKSNRTDVMITAPKNKKMTENLAAPEA LDSSTHSSSTATOSRAKMNTPAPTPSTVPAIPR GGGGPPPCAPHDRVSSVLQCDTQAMDHKTE SSHSVVEFLFKRTKTPSPFHPAVRENN LDSSTHSSSTATOSRAKMNTPAPTPSTVPAIPR GGGGPPPCAPHDRVSSVLQCDTQAMDHKTE SSHSVVEFLFKRTKTPSPFHPAVRENN LGSGGPPPCAPHDRVSSVLQCDTQAMDHKTE SSHSVVEFLFKRTKTPSPFHPAVRENN LGSGLDSCMRTPLCHCPGGPGGELISPCRC	-250	2000	1	1002	1 337	307	ELUAUKSDKEAMEAAVKEEISVEDEAVDKNI
1239 2589 A 9805 105 540 VPGDPAMVRAGAVGAHLPASGLDFGDLKK MNKRQLVYQVLNFAMIVSSALMIWKGLIVT GESPIVVILSGSMEPAFHRGDLLFI.TNFRED PIRAGEIVVFKVEGRDIPIVHRVIKVHEKDNG DIKFLTKGDNHGDDRGSYK DPENDEDHA SEPTIV SE			i				TOEDSTDEOLDE VSKADA VERNOEDSAL
1239 2589 A 9805 105 540 VPODPAMVRAGAVGAHLPASGLDIFGDLIKK MNKRQLYQVLNSGMBPB AHRGDLIFLTHRED PIRAGEIVYFKVEGRDIPIVHEKUNFRED PIRAGEIVYFKVEGRDIPIVHEKUNFRED PIRAGEIVYFKVEGRDIPIVHEKUNFRED PIRAGEIVYFKVEGRDIPIVHEKUNFRED PIRAGEIVYFKVEGRDIPIVHEKUNFRED PIRAGEIVYFKVEGRDIPIVHEKUNFRED PIRAGEIVYFKVEGRDIPIVHEKUNFRED PIRAGEIVYFKVEGRDIPIVHEKUNFRED PIRAGEIVYFKVEGRDIPIVHEKUNFRED PIRAGEIVARARRIGGGECCGAGWVAEWS POPLDPAMIL WMQGFVLEAVACQDNDDYLR YGILFEDLDCROGOVVDIELQEGLRNWSSAF DPNSEEHG							DIEETDI KAREKURNII TEKRETINII 19090RE
1240 2590 A 9805 105				i	ļ		
1240 2590 A 9819 3 305 TORROPHICA STRUCTURE OF PREACTIVETY CEORDIPIVIR VIKY HEXDING DIKFLITK GEORGE CCGAGWYAEWS POPLDP AMPLIL WIND GEORGE CCGAGWYAEWS POPLD PAMPLIL WIND GEORGE CCGAGWYAEWS POPLD PAMPLIL WIND GEORGE CCGAGWYAEWS POPLD PAMPLIL WIND GEORGE CCGAGWYAEWS POPLD PAMPLIL WIND GEORGE CCGAGWYAEWS POPLD PAMPLIL WIND GEORGE CCGAGWYAEWS POPLD PAMPLIL WIND GEORGE CCGAGWYAEWS POPLO PAMPLIL WIND GEORGE CCGAGWYAEWS POPLD PAMPLIL WIND GEORGE CCGAGWYAEWS POPLO PAMPLIL WIND GEORGE CCGAGWYAEWS POPLO PAMPLIL WIND GEORGE CCGAGWYAEWS POPLO PAMPLIL WIND GEORGE CCGAGWYAEWS POPLO PAMPLIL WIND GEORGE CCGAGWYAEWS POPLO PAMPLIL WIND GEORGE CCGAGWY GEORGE CCGAGWYAEWS POPLO PAMPLIL WIND GEORGE CCGAGWYAEWS POPLO PAMPLIL WIND GEORGE CCGAGWYAEWS POPLO PAMPLIL WIND GEORGE CCGAGWYAEWS POPLO PAMPLIL WIND GEORGE CCGAGWYAEWS POPLO PAMPLIL WIND GEORGE CCGAGWYAEWS POPLO PAMPLIL WIND GEORGE CCGAGWYAEWS POPLO PAMPLIL WIND GEORGE CCGAGWYAEWS POPLO PAMPLIL WIND GEORGE CCGAGWYAEWS POPLO PAMPLIL WIND GEORGE CCGAGWYAEWS POPLO PAMPLIL WIND GEORGE CCGAGWYAEWS POPLO PAMPLIL WIND GEORGE CCGAGWYAEWS POPLO PAMPLIL WIND GEORGE CCGAGWYAEWS PAMPLIC P	1239	2589	Δ	0205	105	540	
1240 2590 A 9819 3 305 TDGRDPLPCARREGUELLENGED TDGRDPLPCARREGUE TDGRDPLPCARREGUE TDGRDPLPCARREGUE TDGRDPLPCARREGUE TDGRDPLPCARREGUE TDGRDPLPCARREGUE TDGRDPLPCARREGUE TDGRDPLPCARREGUE TDGRDPLPCARREGUE TDGRDPLPCARREGUE TDGRDPLPCARREGUE TDGRDPLPCARREGUE TDGRDPLPCARREGUE TDGRDPLPCARREGUE TDGRDPCARREG	1237	2307	^	7005	103	1 340	VPGDPAMVKAGAVGAHLPASGLDIFGDLKK
PIRAGEIVYFKVEGRDIPIVIRIVIKYIEKDNG DIKFLTKGDNRGDIDRSGDIRGSDRG DIKFLTKGDNRGDIDRSGDIRGS DIKFLTKGDNRGDIDRSGDIRGS DIKFLTKGDNRGDIDRSGDIRGS DIKFLTKGDNRGDIDRSGDIRGS DIKFLTKGDNRGDIRGS PQRLPPAMLL WMQGFVLEAVACQDNDIV JLR PQRLPPAMLL WMQGFVLEAVACQDNDIV JLR PQRLPPAMLL WMQGFVLEAVACQDNDIV JLR PQRLPPAMLL WMQGFVLEAVACQDNDIV JLR PQRLPPAMLL WMQGFVLEAVACQDNDIV JLR PQRLPPAMLL SKRINKTD MITTAPINKKMTENLAAPEA LDSSTHSSSTATQSRAKMNIPAPIPSTVPAIPR GSGGGPPACPAPIBNVSSVLQCTOTQAMDHKTE SSHSVVEFLFKRTKTPSPFHPAVRENRN TISCGPATEPPASLLSSASSDDIPCKEKTEDRYS LGSSLDSGMRTFLCRICFQCPEQGELLSPCRC DGSVKCTHQPCLIK WISERGCWSCELCYYKY HVIAISTKINPLQWQAISLTKEQVAVALIGS LFILASISWLIWSTFSPSARWQRQDLLFQICVG MYGFMDVMIVAVDSEDMVQAAKLEVGKRWS DIPP DGFHSNFILANAQVAKGFPIVYCSDGFCELAG FARTEVMQ DGFHSNFILANAQVAKGFPIVYCSDGFCELAG FARTEVMQ FARTEVMQ PAGGAGGPGGASARKLSTFLGVVPTVLSMF SIVVFLRIGFVVGHAGLLQALAMLLVAYFILA LTVLSVCAIATNGAVQGGAVCLIQHRWTG VWPVLPAREVMISRTLGPEVGGSIGMFVLA LTVLSVCAIATNGAVQGGAVCLIQHRWTG VWPVLPAREVMISRTLGPEVGGSIGMFVLA LTVLSVCAIATNGAVQGGAVCLIQHRWTG VWPVLPAREVMISRTLGPEVGGSIGMFVLA NVCGCAVSLIGL VESVLDVFGA AMSLTEDVLAAALADHLPEDKWSASKRRPL KSSLGYEITSLLIPPDRSSISSLIPDDMM SYIGFKRTAVVRGIMHREAFNIIGRRIVQVAQ AMSLTEDVLAAALADHLPEDKWSASKRRPL KSSLGYEITSLLIPPDRSSISSLIPDDMM SYIGFKRTAVVRGIMHREAFNIIGRRIVQVAQ AMSLTEDVLAAALADHLPEDKWSASKRRPL KSSLGYEITSLLIPPDRSSISSLIPDDMM SYLGFKRTAVVRGIMHREAFNIIGRRIVQVAQ AMSLTEDVLAAALADHLPEDKWSASKRRPL KSSLGYEITSLLIPPDRSGINSSILPDDMM SYLGFKRTAVVRGIMHREAFNIIGRRIVQVAQ AMSLTEDVLAAAALADHLPEDKWSASKRRPL KSSLGYEITSELNIPPDRSGINSSILPDDMM SYLGFKRTAVVRGIMHREAFNIIGRRIVQVAQ AMSLTEDVLAAAALADHLPEDKWSASKRRPL KSSLGYEITSELNIPPDRSGINSSILPDMM SYLGFKRTAVRGIMYNVORYASKRRPL KSSLGYEITSEGLISPLYIQOKOGAPVAT NAPHSPRWGGMYNNDSKTYAASVLPPVV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LGGPTSEGLIMTWELDRLLWARSVENLATATT TITLISLA CARREPAGADSLSWGAGPRISSV PROSASSSYLDAMSCHARSVENLATATT TITLISLA PROGAGGARSISVDGAGPRISSV PROSASSSYLDAMSCHARSVENLATATT TITLISLA PROGAGGARSISVDGAGPRISSV PROSASSSYLDAMSCHARSVENLA		ľ	1				CSTSPIANT SCOVED A FIX COVED TO THE COVED TO
1240 2590 A 9819 3 305 TDGRDPLPCAARRIGGGECGAGGWVAEWS PQPLDP AMIL WMQGFVLEAVACQDNDD YLR YGILFEDLDCNGGGVTONIELQEGLRNWSSAF PQPLDP AMIL WMQGFVLEAVACQDNDD YLR YGILFEDLDCNGGGVVDIIELQEGLRNWSSAF PQPLDP AMIL WMQGFVLEAVACQDNDD YLR YGILFEDLDCNGGGVVDIIELQEGLRNWSSAF PQPLDP AMIL WMQGFVLEAVACQDNDD YLR YGILFEDLDCNGGGVVDIIELQEGLRNWSSAF PQPLDP AMIL SSTATOSKAKMIPAPPSTYPAPPA GGSGGPPPCAPHDRVSSVLQCDTQAMDHKTE SSHSVVEFFKRTKTPSPFHPAVRNRN GGSGGPPPCAPHDRVSSVLQCDTQAMDHKTE SSHSVVEFFKRTKTPSPFHPAVRNRN LGSSLDSGMRTPLCRICFQGPEGGLISPCRC DGSVKCTHQPCLIKW SSASSDDFCKEKTEDRYS LGSSLDSGMRTPLCRICFQGPEGGLISPCRC DGSVKCTHQPCLIKW SEGCWSCELCYKY HVIAISTKNPLQWQAISLTVIEK VQWAAAILGS LFLIASISWLIWSTTSPSARWQRQDLIFQICYQ MYGFMD WINVAVDSEDMVQAAKEYGKRWS DIPP	1		l				GSESPIVVVLSGSMEPAFHRGDLLFLTNFRED
1240	l		ł				
1241 2591 A 9834 841 1209 SPARGKSNRTDVMITAPKNKKMTENLAAPEA DPNSEEHG SPARGKSNRTDVMITAPKNKKMTENLAAPEA LDSSTHSSSTATQSRAKMNTPAPTSTVPAIPR GSGGPPPCAPHDRVSSVLQCDTQAMDHKTE SSHSVVEFLFKRTKTPSPHHPAVENRN SSHSVVEFLFKRTKTPSPHHPAVENRN LGSSLDSGMRTPLCRICFQGPEQGELLSPCRC DGSVKCTHQPCLIKWISERGCWSCELCYYKY HVIAISTKSPLQWAJAILGS LFILIASISWLIWSTFSPSARWQRQDLLFQICVG MYGFMDVMIVAVDSEDMVQAAKEVGKRWS DIPP DGTHSINFILANAQVAKGFPIVYCSDGFCELAG FARTEVMQ FARTEVMQ FARTEVMQ TGTHSINFILANAQVAKGFPIVYCSDGFCELAG FARTEVMQ TGTHSINFILANAQVAKGFPIVYCSDGFCELAG TGTHSINFILANAQVAKGFPIVTSDGFCENSTVV TGTHSINFILANAQVAKGFPIVTSDGFCENSTVV TGTHSINFILANAQVAKVARVAKOPIPTSTATA TLTSLA TGTHSINFILANAQVAKVARVAKOPTSSKKIWIFGF TGTHSINFILANAQVAKVARVAKVARVARVARVAKVARVAKOPTSKKIWIFGF TGTHSINFILANAQVAKVARVAKVARVAKOPTSKKIWIFGF	1240	2500		0010		205	DIKFLIKGDNNEGDDRGSYK
1241 2591 A 9834 841 1209 SPARGKSNRTDVMITAPKNKKMTENLAAPEA LDSSTHSSTATQSRAKMNTPAAPTPSTVPAIPR GSGGPPCAPHDRYSSVLQCDTQAMDHIKTE SSHSVVEFLFKRTKTPSPFHPAVRENRN I SSHSVVEFLFKRTKTPSPFHPAVRENRN I SSHSVVEFLFKRTKTPSPFHPAVRENRN LGSSLDFCKEKTEDRYS LGSSLDSGMRTPLCRICCQGPEQGELLSPCRC DGSVKCTHQPCLIKWISERGCWSCELCYYKY HVIAISTKSPLQWQAISLTVIEKVQVAAAILGS LFILASISWILWSTFSPSARWQRQDLLFQICVG MYGFMDVMIVAVDSEDMVQAAKEVGKRWS DIPP 1243 2593 A 9846 198 411 WIGSHFJAGKMPYMKGLLAPQNTPLDFHATRF DGTHISNTRLANQQVAKGFPIVYCSDGFCELAG FARTEVMQ PICGFLYLCSAMASESSPLLAYRILGEGVAL PANGAGGPGGASARKLSTFLGVVVPTVLSMF SIVVFLRIGFVVGHAGILQALAMLLVAYFILA LTVLSVCAIATNGAVGGAYCTLQFR WTG VWPVLPAREVMISRTLGPEVGGSIGLMFYLA NVCGCAVSLLGLVESVLDVFGG AYCILQFR WTG VWPVLPAREVMISRTLGPEVGGSIGLMFYLA NVCGCAVSLLGLVESVLDVFGG AYCILQFR WTG VWPVLPAREVALAALADHLPEKWSAEKRPL KSSLGYEITFSLLNPDPKSHDVTWDIEGAVRR YVQPFLAMAGAGAFFARFANIGGRIVQVAQ AMSLTEDVLAAALADHLPEKWSAEKRPL KSSLGYEITFSLLNPDPKSHDVTWDIEGAVRR YVQPFLAMAGAGAFFAPNIGRRIVQVAQ AMSLTEDVLAAALADHLPEKWSAEKRPL KSSLGYEITFSLLNPDPKSHDVTWDIEGAVRR YVQPFLAMAGAGAFFAPNIGRRIVQVAQ AMSLTEDVLAAALADHLPEKWSAEKRPL KSSLGYEITFSLLNPDPKSHDVTWDIEGAVRR YVQPFLAMAGAGAFFAPNIGRRIVQVAQ AMSLTEDVLAAALADHLPEKWSAEKRPL KSSLGYEITFSLLNPDPKSHDVTWDIEGAVRR YVQPFLAMAGAGAFFAPNIGRRIVQVAQ AMSLTEDVLAAALADHLPEKWSAEKRPL KSSLGYEITFSLLNPDPKSHDVTWDIEGAVRR YVQPFLAMAGAGAFFAPNIGRRIVQVAQ AMSLTEDVLAAALADHLPEKWSAEKRPL KSSLGYEITFSLLNPDPKSHDVTWDIEGAVRR YVQPFLAMAGAGAFFAPNIGRRIVQVAQ AMSLTEDVLAAALADHLPEKWSAEKRPL KSSLGYEITFSLLNPDPKSHDVTWDIEGAVRR YVQPFLAMAGAGAFFAPNIGRRIVQVAQ AMSLTEDVLAAALADHLPEKWSAEKRPL KSSLGYEITFSLLNPDPKSHDVTWDIEGAVRR YVQPFLAMAGAGAFFAPNIGRRIVQVAQ AMSLTEDVLAAAALADHLPEKWSAEKRPL KSSLGYEITFSLAPPROBORYVADAAT NAFHSPRWGGIMVVNVDSKTYNASVLPVRV EVDMVRWMEVFLAQLRLLFGLAGPQLPPFCL LSGFFFTTLVPSTRMALGTLALVLALPCK RREPAGADALSWGAGADSLSWGAFFRPSTSVV FRYGKRKKKGAPPSKKIWIPGP GALCSKHRQESDFESVFAKKKKGAPPSSKWIWFGP GALCSKHRQESDFESVFAKKKKGAPPSVS GARLCKHRGRYPSSKIWIPGP GALCSKHRQESDFESVGRIKKKGAPPSVS	1240	2390	A	9019	3	303	TDGRDPLPCAARRRGGGGECCGAGWVAEWS
1241 2591		ŀ					PQPLDPAMLLWMQGFVLEAVACQDNDDYLR
1241 2591			1		!		
1242 2592 A 9843 3 589	1241	2501	Α -	0024	041	1000	
1242 2592 A 9843 3 589	1241	2391	A	9834	841	1209	
1242 2592 A 9843 3 589			ļ				
1242	1		1	[
1243 2593 A 9846 198 411 WRISHHAGKMPVMKGLLAPQNTFLDTHAFRF DGTHSNFLANAULASSAMAKEVGKRWS DIPP 1244 2594 A 9848 116 650 PICGFLYLCSAMASESSPLLAYRLLGEEGVAL PANGAGPGGASARKLSTFLGVVVPTVLSMF SIVVFLRIGFVVGHAGILQALAMLLVAYFILA LTVLSVCALATIAGPEVGGSIGMFYLA NVCGCAVSLLGLVESVLDVFGA AM.DEPQEQAGGAYCILQHRWTG VWPVLPAREVMSRTLGPEVGGSIGMFYLA NVCGCAVSLLGLVESVLDVFGA AM.DEPQEQAGGSTRUTVVISHSHSSLLPQDMM SYIGPKRAVVRGIMHREAFNIIGRRIVQVAQ AMSLTEDVLAAALADHLPEDKWSAEKRPL KSSLGYEIFFSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYVLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYPELAHSPLVIQDKGGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRIPDVRAPLRVTSPGLRSRS ARSLGRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RERPRAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF	1242	2502	<u> </u>	0040			
DGSVKCTHOPCLIKWISERGCWSCELCYYKY HVIAISTKNPLQWQAISLTVIEKVQVAAAILGS LFLIAISISWIJWSTFSPSARWQRQDLLFQICYG MYGFMDVMIVAVDSEDMVQAAKEVGKRWS DIPP 1243 2593 A 9846 198 411 WRISHHAGKMPYMKGLLAPQNTFLDTHATRF- DGTHSNFILANAQVAKGFPIVYCSDGFCELAG FARTEVMQ 1244 2594 A 9848 116 650 PICGFLYLCSAMASESSPLLAYRLLGEEGVAL PANGAGPGGASARKLSTFLGVVVPTVLSMF SIVVFLRIGFVVGHAGLLQALAMLLVAYFILA LTVLSVCAIATNGAVQGGGAYCILQHRWTG VWPVLPAREVMISSRTLGPEVGGSIGLMFYLA NVCGCAVSLLGLVESVLDVFGA 1245 2595 A 9849 573 1620 KSKCRFPEGLSEGFGMRKEALSGSVQEAE AMLDEPGQAEGSLTVYVISEHSSLLPQDMM SYIGPKRTAVVRGIMHREAFNIIGRRIVQVAQ AMSLTEDVLAAALADHLPEDKWSAEKRRPL KSSLGYEITSLNPPPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAOPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFTTVVSTERALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQPFTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS	1242	2392	A	9843	3	589	TISCGPATEPPASLLSSASSDDFCKEKTEDRYS
HVIAISTKNPLQWQAISLTVIEK VQVAAALIGS LFLIASISWLIWSTFSPSARWQRQDLLFQICYG MYGFMDVMIVAVDSEDMVQAAKEVGKRWS DIPP 1243 2593 A 9846 198 411 WRISHHAGKMPVMKGLLAPQNTPLDTIAFRF DGTHSNFILANAQVAKGFPIVYCSDGFCELAG FARTEVMQ 1244 2594 A 9848 116 650 PICGFLYLCSAMASESSPLLAYRLLGEEGVAL PANGAGGPGGASARKLSTFLGVVVPTVLSMF SIVVFLRIGFVVGHAGLLQALAMLLVAYFILA LTVLSVCAIATNGAVQGGGAYCILQHR WTG VWPVLPARWMSRTLGPEVGGSIGLMFYLA NVCGCAVSLLGLVESVLDVFGA 1245 2595 A 9849 573 1620 KSKCRFPEGLSEGFGPMRKEALSSGSVQEAE AMLDEPQEQAEGSLTVYVISEHSSLLPQDMM SYLGPKATAVVRGIMHREAFNIIGRRIVQVAQ AMSLTEDVLAAALADHLPEDKWSAEKRPL KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNIFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLVPLELAFSPLYIQDKDGAPVAT NAPHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFTPLVPSTRAALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS			İ				LGSSLDSGMRTPLCRICFQGPEQGELLSPCRC
LFLIASISWLIWSTFSPSARWQRQDLLFQICYG MYGFMDVMIVAVDSEDMVQAAKEVGKRWS DIPP 1243 2593 A 9846 198 411 WRISHHAGKMPVMKGLLAPQNTFLDTIATRFDGTHSNFILANAQVAKGFPIVYCSDGFCELAG FARTEVMQ 1244 2594 A 9848 116 650 PICGFLYLCSAMASESSPLLAYRLLGEEGVAL PANGAGGPGGASARKLSTFLGVVVPTVLSMF SIVVFLRIGFVVGHAGLLQALAMLLVAYFILA LTVLSVCAIATNGAVQGGGAYCLLQHR WTG VWPVLPAREVMISRTLGPEVGGSIGLMFYLA NVCGCAVSLLGLVESVLDVFGA 1245 2595 A 9849 573 1620 KSKCRFPEGLSEGFGPMRKEALSSGSVQEAE AMLDEPQEQAEGSLTVYVISEHSSLLPQDMM SYIGPKRTAVVRGIMHREAFNIIGRRIVQVAQ AMSLTEDVLAAALADHLPEDKWSAEKRRPL KSSLGYEITTSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVNFLYVPELALBPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRILFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTILVPSTRMALGTLALVLALPCK RREPPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQPFTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS							
1243 2593 A 9846 198 411 WKISHHAGKMPVMKGLLAPQNTFLDTHATRF- DGTHSNFILANAQVAKGFPIVYCSDGFCELAG FARTEVMQ 1244 2594 A 9848 116 650 PICGFLYLCSAMASESSPLLAYRLLGEEGVAL PANGAGOPGGASARKLSTFLGVVVPTVLSMF SIVVFLRIGFVVGHAGLLQALAMLLVAYFILA LITVLSVCAIATNGAVQGGGAYCLLQHRWTG VWPVLPAREVMISRTLGPEVGGSIGMFYLA NVCGCAVSLLGLVESVLDVFGA 1245 2595 A 9849 573 1620 KSKCRFPEGLSEGFGPMRKEALSSGSVQEAE AMILDEPQEQAEGSLTVYVISEHSSLLPQDMM SYIGPKRAVVRGIMHREAFNIIGRRIVQVAQ AMSLTEDVLAAALADHLPEDKWSAEKRPL KSSLGYEITFSLLNPDFKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVWEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFFTDTIORSK WIRA VNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS	1		ŀ	}			HVIAISTKNPLQWQAISLTVIEKVQVAAAILGS
1243 2593 A 9846 198 411 WISHIAGKMPVMKGLLAPQNITELDTIATREDTH DGTHSNFILANAQVAKGFPIVYCSDGFCELAG FARTEVMQ 1244 2594 A 9848 116 650 PICGFLYLCSAMASESSPLLAYRLLGEGVAL PANGAGGPGGASARKLSTFLGVVVPTVLSMF SIVVFLRIGFVVGHAGLLQALAMLLVAYFILA LTVLSVCAIATNGAVQGGGAYCILQHRWTG VWPVLPAREVMISRTLGPEVGGSIGLMFYLA NVCGCAVSLLGL VESVLDVFGA 1245 2595 A 9849 573 1620 KSKCRFPEGLSEGFGPMRKEALSSGSVQEAE AMI.DEPQEQAEGSLTVYVISEHSSLLPQDMM SYIGPKRTAVVRGIMHREAFNIIGRRIVQVAQ AMSLTEDVL.AAALADHLPEDK WSAEKRRPL KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPYLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSSCFFFTILVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNVDPPSKKLWIPGP GAILCSCHFQESDFESYGIRRKLKKGAVPSVS	1						LFLIASISWLIWSTFSPSARWQRQDLLFQICYG
1243 2593 A 9846 198 411 WRISHFIAGKMFVMKGLLAPQNTFLDTTATRFDGTHSNFILANAQVAKGFPIVYCSDGFCELAG FARTEVMQ 1244 2594 A 9848 116 650 PICGFLYLCSAMASESSPLLAYRLLGEEGVAL PANGAGGPGGASARKLSTFLGVVVPTVLSMF SIVVFLRIGFVVGHAGLLQALAMLLVAYFILA LTVLSVCAIATNGAVQGGGAYCILQHRWTG VPVVLPAREVMISRTLGPEVGGSIGLMFYLA NVCGCAVSLLGLVESVLDVFGA 1245 2595 A 9849 573 1620 KSKCRFPEGLSEGFGPMRKEALSSGSVQEAE AMLDEPQEQAEGSLTVYVISEHSSLLPQDMM SYIGPKRTAVVRGIMHREAFNIIGRRIVQVAQ AMSLTEDVLAAALADHLPEDKWSAEKRPL KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVVV EVDMVVMEVFLAQRALLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKLIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS	1						
1244 2594 A 9848 116 650 PICGFLYLCSAMASESSPLLAYRLLGEGVAL PANGAGGPGGASARKLSTFLGVVPTVLSMF SIVVFLRIGFVVGHAGLLQALAMLLVAYFILA LTVLSVCAIATNGAVQGGAYCILQHRWTG VWPVLPAREVMISRTLGPEVGGSIGLMFYLA NVCGCAVSLLGVESVLDVFGA SKKCRFPEGLSEGFGPMRKEALSSGSVQEAE AMLDEPQEQAEGSLTVYVISEHSSLLPQDMM SYIGPKRTAVVRGIMHREAFNIIGRRIVQVAQ AMSLTEDVLAAALADHLPENKWSAEKRRPL KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRILFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIORSKWIRAVNVDPRSKLKIMIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS	1242	2502	•	0046	100		
1244 2594 A 9848 116 650 PICGFLYLCSAMASESSPLLAYRLLGEEGVAL PANGAGGPGGASARKLSTFLGVVVPTVLSMF SIVVFLRIGFVVGHAGLLQALAMLLVAYFILA LTVLSVCAIATNGAVQGGGAYCILQHRWTG VWPVLPAREVMISRTLGPEVGGSIGLMFYLA NVCGCAVSLLGL VESVLDVFGA 1245 2595 A 9849 573 1620 KSKCRFPEGLSEGFGPMRKEALSSGSVQEAE AMLDEPQEQAEGSLTVYVISEHSSLLPQDMM SYIGPKRTAVVRGIMHREAFNIIGRRIVQVAQ AMSLTEDVLAAALADHLPEDKWSAEKRRPL KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVONYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS	1243	2393	A	9840	198	411	
1244 2594 A 9848 II6 650 PICGFLYLCSAMASESSPLLAYRLLGEEGVAL PANGAGGPGASARKLSTFLGVVVPTVLSMF SIVVFLRIGFVVGHAGLLQALAMLLVAYFILA LTVLSVCAIATNGAVQGGGAYCILQHRWTG VWPVLPAREVMISRTLGPEVGGSIGIMFYLA NVCGCAVSLLGLVESVLDVFGA 1245 2595 A 9849 573 1620 KSKCRFPEGLSEGFGPMRKEALSSGSVQEAE AMLDEPQEQAEGSLTVYVISEHSSLLPQDMM SYIGPKRTAVVRGIMHREAFNIIGRRIVQVAQ AMSLTEDVLAAALADHLPEDKWSAEKRPL KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HOPPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS							
PANGAGGPGASARKLSTFLGVVYPTVLSMF SIVVFLRIGFVVGHAGLLQALAMLLVAYFILA LTVLSVCAIATNGAVQGGAYCILQHRWTG VWPVLPAREVMISRTLGPEVGGSIGLMFYLA NVCGCAVSLLGLVESVLDVFGA 1245 2595 A 9849 573 1620 KSKCRFPEGLSEGFGPMRKEALSSGSVQEAE AMLDEPQEQAEGSLTVYVISEHSSLLPQDMM SYIGPKRTAVVRGIMHREAFNIIGRRIVQVAQ AMSLTEDVLAAALADHI.PEDKWSAEKRRPL KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HOFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS	1244	2504		0040			
SIVVFLRIGFVVGHAGLLQALAMLLVAYFILA LTVLSVCAIATNGAVQGGAYCILQHRWTG VWPVLPAREVMISRTLGPEVGGSIGLMFYLA NVCGCAVSLLGLVESVLDVFGA 1245 2595 A 9849 573 1620 KSKCRFPEGLSEGFGPMRKEALSSGSVQEAE AMLDEPQEQAEGSLTVYVISEHSSLLPQDMM SYIGPKRTAVVRGIMHREAFNIIGRRIVQVAQ AMSLTEDVLAAALADHLPEDKWSAEKRRPL KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TILTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRA VNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS	1244	2394	A	9848	116	650	PICGFLYLCSAMASESSPLLAYRLLGEEGVAL
LTVLSVCAIATNGAVQGGAYCILQHRWTG VWPVLPAREVMISRTLGPEVGGSIGLMFYLA NVCGCAVSLLGLVESVLDVFGA 1245 2595 A 9849 573 1620 KSKCRFPEGLSEGFGPMRKEALSSGSVQEAE AMLDEPQEQAEGSLTVYVISEHSSLLPQDMM SYIGPKRTAVVRGIMHREAFNIIGRRIVQVAQ AMSLTEDVLAAALADHLPEDKWSAEKRPL KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS	ŀ						
VWPVLPAREVMISRTLGPEVGGSIGLMFYLA NVCGCAVSLLGLVESVLDVFGA 1245 2595 A 9849 573 1620 KSKCRFPEGLSEGFGPMRKEALSGSVQEAE AMLDEPQEQAEGSLTVYVISEHSLLPQDMM SYIGPKRTAVVRGIMHREAFNIIGRRIVQVAQ AMSLTEDVLAAALADHLPEDKWSAEKRRPL KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTTQRSK WIRAVNRVDPRSKKTWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS							
1245 2595 A 9849 573 1620 KSKCRFPEGLSEGFGPMRKEALSSGSVQEAE AMILDEPQEQAEGSLTVYVISEHSSLLPQDMM SYIGPKRTAVVRGIMHREAFNIIGRRIVQVAQ AMSLTEDVLAAALADHLPEDKWSAEKRRPL KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS							
1245 2595 A 9849 573 1620 KSKCRFPEGLSEGFGPMRKEALSSGSVQEAE AMLDEPQEQAEGSLTVYVISEHSSLLPQDMM SYIGPKRTAVVRGIMHREAFNIIGRRIVQVAQ AMSLTEDVLAAALADHLPEDKWSAEKRRPL KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS	}						VWPVLPAREVMISRTLGPEVGGSIGLMFYLA
AMLDEPQEQAEGSLTVYVISEHSSLLPQDMM SYIGPKRTAVVRGIMHREAFNIGRRIVQVAQ AMSLTEDVLAAALADHLPEDKWSAEKRPL KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS	1245	2505		0040		1600	
SYIGPKRTAVVRGIMHREAFNIIGRRIVQVAQ AMSLTEDVLAAALADHLPEDKWSAEKRRPL KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS	1245	2293	A	7047	3/3	1020	
AMSLTEDVLAAALADHLPEDKWSAEKRPL KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR YVQFFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS]				J	ŀ	
KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS		ł				ł	
YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS	Į			. [ļ	
RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS	'		İ		ł	l	
SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS							YVQPFLNALGAAGNFSVDSQILYYAMLGVNP
NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS		1					RFDSASSSYYLDMHSLPHVINPVESRLGSSAA
EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV RRERPAGADSLSWGAGPRISSYV FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS	[i	ĺ	1	l		
LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA				1			NAFHSPRWGGIMVYNVDSKTYNASVLPVRV
1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSK WIRA VNRVDPRSKK IWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS							EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL
1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS				- 1	1		
ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS	1345	2507]	0860			
QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS	1240	2390	A	טכאנ	114	464	
1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS				ł	l		
1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS		Ì					
HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS	1245	0.505					
HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS	1247	2597	Α	9851	2	327	FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF
GAILCSKHFQESDFESYGIRRKLKKGAVPSVS		İ	1			Ì	HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP
LYKVFKYSSRCTS	'	J]	J	J	}	GAILCSKHFQESDFESYGIRRKLKKGAVPSVS
							LYKVFKYSSRCTS

NO. of plant of the color of period control of the color of period control of period period period pe	CEO ID	SEQ ID	Mot	LCEA	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucleotide seq- uence 194	SEQ ID		Met	SEQ			
USSN Ocation			liou	1 .			
1248 2598		1	İ				
1248			I.				
amino acid residue of peptide residue of peptide residue of peptide sequence T-Threonine, V-V-Line, W-Tryptophen, Y-Stop oddon, Y-Stop odd	-	ucice		1			
Pesidue of peptide sequence	uence			914			
Poptide sequence		ĺ	1	1		1 .	
1248 2598 A 9853 58 444 RVDPEVYSKGKDAGGADVSLACRROSIPEE FRGITVVELIKEGSTI.GLTISGGTDKDGKPR FRGITVVELIKEGSTI.GLTISGGTDKDGKPR VSNLRPGGLAARSDLINGDYIRSVGKIHLTR LRHDEIITLLKNVGERVVLEVEYELPPPGCP WT LRHDEIITLLKNVGERVVLEVEYELPPPGCP WT LARDEIITLLKNVGERVVLEVEYELPPPGCP WT LARDEIITLLKNVGERVVLEVEYELPPPGCP WT LARDEIITLLKNVGERVVLEVEYELPPPGCP WT LARDEIITLLKNVGERVVLEVEYELPPPGCP WT LARDEIITLKNVGERVVLEVEYELPPPGCP WT LARDEIITLKNVGERVALEVEYELPPPGCP WT RAPVRGQDSGAGTPAGRAAAAAGPTVSAV RAPVRGQDSGAGTPAGRAAAAAGPTVSAV RAPVRGQDSGAGPAGRAGASTAGAGAAAAAGPTVSAV GAGAGAAAAAGPAGRAHIAFRRCADAGEAVGASC GRCAVALLSCVCTLVSTHVCVGSGCPGAAGI PMGAGGAGASASSAVTTAFQEPPARIQAGS GAGPAPGRAMKSTTLLALLALVLLVLVSGAL VFRALEQPHEQQAQREGEVREKFIRAHPCV SSQELGLILIKEVADALGGGAGPETNSTSNSSS SAWDLGSAFFFSGTIUTTIGGGDWINFINGGK ELPHIGGRCRETEGSQVAPRLPASPLCPGYGN VALRTDAGRLFCHTVFVCYMEDWSKLEAY FVIVTLTITVGGDLYAV VALRTDAGRLFCHTVFVCYMEDWSKLEAY FVIVTLTITVGGDLYAV LORLAG WEWLI FWAYAFGGCVHISLTAHIDISHNAAFGTGR LARVALVLVLVVQHALACWLVRGLAWRWLI FWAYAFGGCVHISLTAHIDISHNAAFGTGR ARRWALVLVLVLVQHALACWLVRGLAWRWLI FWAYAFGGCVHISLTAHIDISHNAAFGTGR ARRWALVLVLVLVQHALACWLVRGLAWRWLI FWAYAFGGCVHISLTAHIDISHNAAFGTGR ARRWALVLVLVLVQHALACWLVRGLAWRWLI FWAYAFGGCHAPKACHA WRWIL FWAYAFGGCHAPKACHA WRWIL FWAYAFGGCHAPKACHA WRWIL FWAYAFGGCHAPKACHA WRWIL FWAYAFGGCHAPKACHA WRWIL FWAYAFGGCHAPKACHA WRWIL FWAYAFGGCHAPKACHA WRWIL FWAYAFGGCHAPKACHA WRWIL FWAYAFGGCHAPKACHA WRWIL FWAYAFGGCHAPKACHA WRWIL FWAYAFGGCHAPKACHA WRWIL FWAYAFGCCWISFRCHAPKACHA WRWIL FWAYAFGGCHAPKACHA WRWIL FWAYAFGCCWISFRCHAPKACHA WRWIL FWAYAFGCCWISFRCHAPKACHA WRWIL FWAYAFGCCCWISFRCHAPKACHA WRWIL FWAYAFGCCCWISFRCHAPKACHA WRWIL FWAYAFGCCCWISFRCHAPKACHA WRWIL FWAYAFGCCCWISFRCHAPKACHA WRWIL FWAYAFGCCCWISFRCHAPKACHA WRWIL FWAYAFGCCCWISFRCHAPKACHA WRWIL FWAYAFGCCCWISFRCHAPKACHA WRWIL FWAYAFGCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC			ŀ		4	sequence	
1248 2598		1	1	!	,	1	
PRGITVVELIKEGSTLGLTISGTIDKDGKPR VSNLRPGGLARSDLINGDYIRSVGIBLTR LRHDEIITLLKNVGERVVLEVYELPPPGGP VSNLRPGGLARSDLINGDYIRSVGIBLTR LRHDEIITLKNVGERVVLEVYELPPPGGP WT LRHDEIITLKNVGERVVLEVYELPPPGGP WT LRHDEIITLKNVGERVVLEVYELPPPGGP WT LRHDEIITLKNVGERVVLEVYEYELPPPGGP WT LRHDEIITLKNVGERVLEVYEYELPPPGGP WT LRHDEIITLKNVGERVLEVYEYELPPPGGP WT LRHDEIITLKNVGERVLEVYESGCPGAGT GRGVAARIAFBURRCADAGEAVGASC GRCAVALLSGVCTLVSTHVCVGSGCPGAGT GRGAGDAGASAESAVTTAPQEPPARPLQAGS GAGPAPGRAARISTILLALLALVILVILVSGGAL VFRALEQPHEQQAGREGOVERKETRAHPCV SDQELGLIKEVADALOGGADPETNSTSNSSS SAWDLGSAFFFSGTITTHOGGDWINVGGK ELPHGGRCRETEGSQVAPRLPASPLCPGYGN VALRTDAGRLFGLTYFVCYMEDWSKLEANY FVIVTLTIVGGDVA GRGSSLRHGGHEAFILKWHVPPELVEVLSA M.P.LLIGCLLFVLTPTFVCYMEDWSKLEANY FVIVTLTITVGGDVA GRGSSLRHGGHEAFILKWHVPPELVEVLSA M.P.LLIGCLLFVLTPTFVCYMEDWSKLEANY FVIVTLTITVGGDVA GRGSSLRHGGHEAFILKWHVPPELVEVLSA M.P.LLIGCLLFVLTPTFVCYMEDWSKLEANY FVIVTLTITVGGDVA GRGSSLRHGGHEAFILKWHVPPELVEVLSA M.P.LLIGCLLFVLTPTFVCYMEDWSKLEANY FVIVTLTITVGGDVA GRGSSLRHGGHEAFILKWHVPPELVEVLSA M.P.LLIGCLLFVLTPTFVCYMEDWSKLEANY FVIVTLTITVGGDVA GRGSSLRHGGHEAFILKWHVPPELVEVLSA GRGVAGARNA GRGVA			 				
VSNLRPGGLARSDLLNIGDYIRSVNOIHLTR LRIBBITLLKNIVGERVVLEVEYELPPPGGCP WT	1248	2598	A	9853	58	444	
LRIDBITLLKNVGERVVLEVEYELPPPGGCP WT							
1249			l	1		•	
1249 2599 A 9856 2 1265 LPPPEPERHRGRAGTRASAAAAAGPTVSAV RAPVROQDSGAGTPOGRLAGRALISVOG SGSCVAAGPARILAPRICACO SGSCVAAGPARILAPRICACO SGSCVAAGPARILAPRICACO SGSCVAAGPARILAPRICACO SGSCVAAGPARILAPRICACO SGACPAPGRAMRSTILLALLALLISVICYOGACPAGO PMGAGDAGSASSA VITTAPCPPARPIQACO GACPAPGRAMRSTILLALLALLISVICYOGACPAGO SOLECHLIKEVADALGGADPETNISTSSISSIS SAWDIGSAPPESGTILITIGGGGDWHVGGGK ELPHGGRCRETEGSQVAPRLPASPLCPGYON VALRTDAGRLECIPYALVOGIPELASVICO RLGSSLRIGIGHEAIPLKWHVPPELVRVLSA MLFLLIGCLLFVLTPTFVFCYMEDWSKLEAIY FVVYTLTTVOFGDYVA FVVYPCOGIPGRAPPNGASRPTMGNSASRNDH EWYYTDOPHTQRKELLAKYPARKALMRPPP RLKWAVLVLVQMLACWLVRGLAWR FVVYTLSTVOFGDYVA SARNRWLAVBANLPEGGVPYAASFKKYNDH HRVLGGDGLDVDVPTILEGWPFCTPARKLL WLVLQPFFYSLRPLCVPHFAVTRMEVLNTLV QLA CLA SARNRWLAVBANLPEGGVPYAASFKKYNDH HRVLGGDGLDVDVPTILEGWPFCTPARKLL WLVLQPFFYSLRPLCVPHFAVTRMEVLNTLV QLA LESYRPOTIDLSRDTIGCALQHISDRENIDDIN MEENPSDHPRASTIFLSKSQTDVREKRISJEN HHPPGGIARXVSSCSTIFIDDSTVGPKKRISJEN HHPPGGIARXVSSCSTIFIDDSTVGPKKRISJEN HHPPGGIARXVSSCSTIFIDDSTVGPKKRISJEN HHPPGGIARXVSSCSTIFIDDSTVGPKKRISJEN HHPPGGIARXVSSCSTIFIDDSTVGPKKRYSTEN HHPPGGIARXVSSCSTIFIDDSTVGPKKRYSTEN HHPPGGIARXVSSCSTIFIDDSTVGPKKRYSTEN HHPPGGIARXVSSCSTIFIDDSTVGPKKRYSTEN HHPPGGIARXVSSCSTIFIDDSTVGPKKRYSTEN HHPPGGIARXVSSCSTIFIDDSTVGPKKRYSTEN HHPPGGIARXVSSCSTIFIDDSTVGPKKRYSTEN HHPPGGIARXVSSCSTIFIDDSTVGPKKRYSTEN SKSEVPPDVKHNFEQKQVTKFVRTLFSAAQI TAECANTLIVYLELIPANWKRV LGALLASKVWDDQAVWNDVCQILKDITV DMWELRQFIELLQPNINVRSSVAXYSYPDL RSLABANNISPLEGGELLYWKAT LGALLASKVWDDQAVWNDVCQILKDITV LGALLASKVWDDQAVWNDVCQILKDITV LGALLASKVWDDQAVWNDVCQILKDITV LGALLASKVWDDQAVWNDVCQILKDITV LGALLASKVWDDQAVWNDVCQILKDITV LGALLASKVWDDQAVWNDVCQILKDITV LGALLASKVWDDQAVWNDVCQILKDITV LGALLASKVWDDQAVWNDVCQILKDITV LGALLASKVWDDQAVWNDVCQILKDITV LGALLASKVWDDQAVWNDVCQILKDITV LGALLASKVWDDQAVWNDVCQILKDITV LGALLASKVWDDQAVWNDVCQILKDITV LGALLASKVWDDQAVWNDVCQILKDITV LGALLASKVWDDQAVWNDVCQILKDITV LGALLASKVWDDQAVWNDVCQILKDITV LGALLASKVWDDQAVWNDVCQILKD		1		1			
RAPVRGQDSGAGTPQGRLAGRGAHLSRVG SGSGVAGPAARHAPRRRCADAGEAVGAS SGSGVAGPAARHAPRRRCADAGEAVGAS GRCAVALLSGVCTLVSTHVCVGSGCPGAAGT PMGAGDAGASABSAVTTAPQEPPARPLQAG GRCAVALLSGVCTLVSTHVCVGSGCPGAAGT PMGAGDAGASABSAVTTAPQEPPARPLQAG GRCAVALLSGVCTLVSTHVCVGSGCPGAAGT PMGAGDAGASABSAVTTAPQEPPARPLQAG SQCAPAGRAMRSTTLLALLALVLLVSLVSGAL VFRALEQPHEQQAGRELGEVREKFLRAHPCV SQCELGLIKEVADALAGGGADPTSTSNSSS SAWDLGSAFFFSGTIITTIGGGDWHVGGGG ELPHGGRCRETEGSQVAPELASPLCPGYGN VALRTDAGRLPCIFYALVGIPLEGILLAGVOD RLGSSLRRIGGHEAFILEWHVPSTLVAVISA MLFLLIGCLLFVLTPTFVFCYMEDWSKLEAIY FVVTTLTTVGFGDYVA ALLAGRAVLVLVLVVQMLACWLVGAWRWLL FWAYAFGGCVNHSLTLAHDISHNAAFGTTR AARNRWLAVFANLYOMALAWRWLL FWAYAFGGCVNHSLTLAHDISHNAAFGTTR AARNRWLAVFANLYOMALAWRWLL FWAYAFGGCVHSLTRAHDISHNAAFGTTR AARNRWLAVFANLPEGVPYAASFKKYHVDH HRYLGGDGLDVDVPTRLEGWFFCTPARKLL WLVLQPFFYSLRPLCVHPKAVTRMEVLNTLV QLA PVIMPLHFSSFGDIVGPSCCVSSSPKLERNAHSR LESYAPDTDLSREDITGCNLQHISDRRINDDLN MERNESDHFRASTITLSKSQTDVREKRISLFIN HHPPGGIARRYSSCSTIFLDDSTVARYSTYL LGALLASKVWDDGAWNAVYOQULKDITVE DMSELERQPLELLQFNINVPSSVYAKYYFOLL GALLASKVWDDQAVWNVDVQULKDITVE DMSELERQPLELLQFNINVPSSVYAKYYFOLL GALLASKVWDDQAVWNVDVQULKDITVE DMSELERQPLELLQFNINVPSSVYAKYYFOLL GALLASKVWDDQAVWNVDVQULKDITVE DMSELERQPLELLQFNINVPSSVYAKYYFOLL GALLASKVWDDQAVWNVDVQULKDITVE DMSELSRQPLELLQFNINVPSSVYAKYYFOLL GRANALSFREDITSRESSENSASKLILLSLYQ VISPGITDTIMDSSTAHSPVLFUPPELSSPSVPAMSC VISPGITDTIMDSSTAHSPVLFUPPELTASEYE STELSATTFSTQSPLQRLFARKMKILGTIQLIF GIMTESFGVIFLFTLLKPYPRFFIFLGVDYPP STELSATTFSTQSPLQRLFARKMKILGTIQLIF GIMTESFGVIFLFTLLKPYPRFFIFLGVDYPP GALAGIILLTFEHPRSKLILL SRIPTSTQSPLQRLFARKMKILGTIQLIF GIMTESFGVIFLFTLLKPYPRFFIFLGVDYPP SGGPAGLLHRPVLFPRMGSCALLIFLNYLOGNQ SVLFRNSGAFLAVKRCTTETLILISRIMMYENSA LGALAGIILLTFEHPRSKLILL LGALAGIILLTFEHPRSKLILL THE PROPERT OF THE PROPERT OF THE PROPERT OF THE PROPERT OF THE PROPERT OF THE PROPERT OF THE PROPERT OF THE PROPERT OF THE PROPERT OF THE PROPERT OF THE PROPERT OF THE PROPERT OF THE PROPERT OF THE PROPERT OF THE PROPERT OF T		L	Í	1		Í	WT
SGSCWAAGPARHIAPRRRCADAGEAVGGAGG GRCAVALLSGVCTU-STHVCVGSGCPGAAGG GRCAVALLSGVCTU-STHVCVGSGCPGAAGG PMGAGDAGASAESAVTTAPQEPPARPLQAGS GAGPAPGRAMRSTITLALLALULI/LYCSGCPGAAGG CAGPAPGRAMRSTITLALLALULI/LYCSGCPGAAGG APBEGAMRSTITLALLALULI/LYCSGCPGAAGG VFRALEQPHEQQACRELGEVREKFLRAHPCV SDQELGLLIKEVADALGGGADPETNSTSNSSS SAWDUGSAPFFSGTITTIGGGGDDWHVGGGK ELPHGGRCRETEGSQVAPRLPASPLCPGYGGV VALRTDAGRLECIFYALVGPLEGILLAGVGD RLGSSLRIGIGHIEAIPLLWHVPPELVRVLSA MLFLLIGCLLFVLTPTFVFCYMEDWSKLEAIY FVIVILITIVGFGDYVA MLFLLIGCLLFVLTPTFVFCYMEDWSKLEAIY FVIVILITIVGFGDYVA MLFLLIGCLLFVLTPTFVFCYMEDWSKLEAIY FVIVILITIVGFGDYVA MLFLLIGCLLFVLTPTFVFCYMEDWSKLEAIY FVIVILITIVGFGDYVA MLFLLIGCLLFVLTPTFVFCYMEDWSKLEAIY FVIVILITIVGFGDYVA FVYPSPCGIPGRAPNGSSRPTMGNSASRNDI FWYPSPGGPGRAFNGASRPTMGNSASRNDI FWYPSPGGPGRAFNGASRRTHAGNSASRNDI FWYPSPGGPGRAFNGASRRTHAGNSASRNDI FWYPSPGRAFICHACMUSTICALULIFYACHTALLAGUSCHAPHACHACHACHACHACHACHACHACHACHACHACHACHACH	1249	2599	Α	9856	2	1265	LPPPRPSRHRRGRAGTRASAAAAAGPTVSAV
GRCAVALLSGVCTLVSTHVCVGSGCPGAACI PMGAGDAGASASASAVTTAPQPARIO_AGS GAGPAPGRAMRSTTLLALLALVLLYLVSGAL VFRALEQPHEQQAGRELGEVREKFLARHPCV SOQELGLIKEVADALGGGADPTSTSNSS SAWDLGSAFFFSGTIITTIGGGDWHVGGGE ELPHGGRCRETEGSQVAPELPASPLCPGYGN VALRTDAGRLFCIFYALVGIPLFGILLAGVGD RLGSSLRRIGGHEAFILEKWHVPSTLVVLSA MLFLLIGGLLFVLTPIFYFCYMBDWSKLEAIY FVLVTLTTVGFGDYVA FVLVTLTTVGFGDYVA FVLVTLTTVGFGDYVA FVLVTLTTVGFGDYVA FVLVTLTTVGFGGDYVA FVLVTLTTVGFGGDYVA FVLVTLTTVGFGGDYVA FVLVTLTTVGFGGLAKYPAIKALMRPDF RLKWAVLVLVLVQMLACWLVGAWRWLL FWAYAFGGCVNISLTLAHIDISHNAAFGTT AARNWLAVFANLPEGVFYAASFKKYHVDH HRYLGGDGLDVDVFTRLEGWFFCTFARKLL WLVLQFFFYSLERJCVHPKAVTTVKVINTLV QLA FVLYPTDLSREDTGCNLQHISDRENIDDLN MERNFSDIFRASTIFLSKSQTDVREKRSI, FNL HRYLGGDGLDVDVFTRLEGWFFCTFARKLL WLVLQFFFYSLERJCVHPKAVTTVSQPNLKYTI KCVALAIYYHIKNRDPDGRMLLDIFDENLHPL SKSEVPPDYKHNPEQKQTRFYRTLFSAAQL TAECAIVTLYYLERLLTYAEIDICPANWKRT LGAILLASKWMDDQAVWNDVOQULKDITVE DMNELERQFLELLQFNINVPSSVYAKYYPI BRLABANNLSFPLEJSREPAHKLEAISRICED KYKDLRRSARKRSASADNLTLPRWSPAIIS KRPDSRPLEJSREPAHKLEAISRICED KYKDLRRSARKRSASADNLTLPRWSPAIIS KRPDSRPLEJSREPAHKLEAISRICED KYKDLRRSARKRSASADNLTLPRWSPAIIS KRPDSRPLEJSREPAHKLEAISRICED KYKDLRRSARKRSASADNLTLPRWSPAIIS KRPDSRPLEJSREPAHKLEAISRICED KYKDLRRSARKRSASADNLTLPRWSPAIIS 1252 2602 A 9879 6 376 KRPDSRPLSCRSGSSKLILSLYQ WINSSEVCSEPRSJSCSRSSKLILSLYQ WINSSEVCSEPRSJSCSRSSKLILSLYQ WINSSEVCSEPRSJSCSRSSKLILSLYQ WINSSEVCSEPRSJSCSRSSKLILSLYQ WINSSEVCSEPRSJSCSRSSSKLILSLYQ WINSSEVCSEPRSJSCSRSSSKLILSLYQ WINSSEVCSEPRSJSCSRSSSKLILSLYQ WINSSEVCSEPRSJSCSRSSKLILSLYQ WINSSEVCSEPRSJSCSRSSKLILSLYQ WINSSEVCSEPRSJSCSRSSKLILSLYQ WINSSEVCSEPRSJSCSRSSKLILSLYQ WINSSEVCSEPRSJSCSRSSSKLILSLYQ WINSSEVCSEPRSJSCSRSSSKLILSLYQ WINSSEVCSEPRSJSCSRSSSKLILSLYQ WINSSEVCSEPRSJSCSRSSKLILSLYQ WINSSEVCSEPRSJSCSRSSKLILSLYQ WINSSEVCSEPRSJSCSRSSSKAN WINSSEVCSEPRSJSCSRSSSKAN WINSSEVCSEPRSJSCS		ļ		i			RAPVRGQDSGAGTPQGRLAGRGAHLSRVGA
PMGAGDAGASAESAVTTAPQEPPARPI,QAGS GAGPAPGRAMSTITLLAILVILVISVGAL VFRALEQPHEQQAQRELGEVREKFLRAHPCV SDQELGILIKEVADALGGADPETNSTNSSS SAWDLGSAFPESGTILTTIGGGGWHVGGGK ELPHGGRCRETEGSQVAPRLPASPL-CPGYGN VALRETDAGRIF-GEFVALVGPLFGILLAGVGD RLGSSLRHGIGHIEAIFLK-WHVPPELVRVLSA MLFLLIGCLL-PUT-PTFVEVYMEDWSKLEAIY FVIVTLTTVGFGDYVA PVIVTLTTVGFGDYVA PVIVTLTTVGFGDYVA PVIVTLTTVGFGDYVA PVIVTLTTVGFGDYVA PVIVTLTTVGFGDYVA PVIVTLTTVGFGDYVA PVIVTLTTVGFGDYVA PVIVTLTTVGFGDYVA PVIVTLTTVGFGDYVA PVIVTLTTVGFGDYVA PVIVTLTTVGFGDYVA PVIVTLTVGFGDYVA PVIVTLLV, VQMLACWLVRGLAKWRUL PWAYAFGGCVMHSLTLAIHDISINAAFGTGF AARNWLAVFANL-PGGVPYAASFKKYHVDH HRYLGGGLDUDVPTRLEGWFCTPARKLLL WLVLQPFFYSLRPLCVHPKAVTRMEVLNTLV QLA ARNRWLAVFANL-PGGVPYAASFKKYHVDH HRYLGGGLDUDVPTRLEGWFCTPARKLLL WLVLQPFFYSLRPLCVHPKAVTRMEVLNTLV QLA LESYRPDTDLSREDTGCNLQHISDRENDDLN MEFNPSDHPRASTIFLJKSGTDVREKKRSLFIN HHPPGQLARKYSSCSTIFLDDSTVSQPNLKYTI KCVALAYYHIKNRDPDGRMLDDFEDNLHPH SKSEVPPDVDKINFPBGKQNTRFVRTLFSAAQL TAECAIVTLVYLERLLTYAEIDICPANWKRIV LIGAILLASK VWDDQAVWNVPYCQILKDITVE DMNELERQFLELLQFNINVPSSVYAKYYDL RSLAEANNLSPHLEPLSRRFAHKLEAISRICED KYKDLRRSARKRSASADNLTLPRWPSDIS KYRDLRRSARKRSASADNLTLPRWPSDIS KYRDLRRSARKRSASADNLTLPRWPSDIS KYRDLRRSARKRSASADNLTLPRWPSDIS KYRDLRRSARKRSASADNLTLPRWPSDIS KYRDLRRSARKRSASADNLTLPRWPSDIS KYRDLRRSARKRSASADNLTLPRWPSDIS KYRDLRRSARKRSASADNLTLPRWPSDIS KYRDLRRSARKRSASADNLTLPRWPSDIS KYRDLRRSARKRSASADNLTLPRWPSDIS KYRDLRSARKRSASADNLTLPRWPSDIS KYRDLRSARKRSASADNLTLPRWPSDIS KYRDLRSARKRSASADNLTLPRWPSDIS KYRDLRSARKRSASADNLTLPRWPSDIS KYRDLRSARKRSASADNLTLPRWPSDIS KYRDLRSARKRSASADNLTLPRWPSDIS KYRDLRSARKRSASADNLTLPRWPSDIS KYRDLRSARKRSASADNLTLPRWPSDIS KYRDLRSARKRSASADNLTLPRWPSDIS KYRDLRSARKRSASADNLTLPRWPSDIS KYRDLRSARKRSASADNLTLPRWPSDIS KYRDLRSARKRSASADNLTLPRWPSDIS KYRDLRSARKRSASADNLTLPRWGFELLYWKAT KAVGIKMGSLSTANVEFCLDVFRELISKNYA KAVGIKMGSLSTANVEFCLDVFRELSKNYA KAVGIKMGSLSTANVEFCLDVFRELSKNYA KAVGIKMGSLSTANVEFCLDVFRELSKNYA KAVGIKMGSLSTANVEFCLDVFRELSKNYA KAVGIKMGSLSTANVEFCLDVFRELSKNYA KAVGIKMGSLSTANVEFCLDVFRELSKNYA KAVGIKMGSLSTANVEFCLDVFRELSKNYA KAVGIKMGSLSTANVEFCLDVFRELSKNYA KAVGIKMGSLSTANVEFC			1	1		ĺ	SGSGVAAGPAARHAPRRRCADAGEAVGASC
PMGAGDAGASAESAVTTAPQEPPARPI,QAGS GAGPAPGRAMSTITLLAILVILVISVGAL VFRALEQPHEQQAQRELGEVREKFLRAHPCV SDQELGILIKEVADALGGADPETNSTNSSS SAWDLGSAFPESGTILTTIGGGGWHVGGGK ELPHGGRCRETEGSQVAPRLPASPL-CPGYGN VALRETDAGRIF-GEFVALVGPLFGILLAGVGD RLGSSLRHGIGHIEAIFLK-WHVPPELVRVLSA MLFLLIGCLL-PUT-PTFVEVYMEDWSKLEAIY FVIVTLTTVGFGDYVA PVIVTLTTVGFGDYVA PVIVTLTTVGFGDYVA PVIVTLTTVGFGDYVA PVIVTLTTVGFGDYVA PVIVTLTTVGFGDYVA PVIVTLTTVGFGDYVA PVIVTLTTVGFGDYVA PVIVTLTTVGFGDYVA PVIVTLTTVGFGDYVA PVIVTLTTVGFGDYVA PVIVTLTTVGFGDYVA PVIVTLTVGFGDYVA PVIVTLLV, VQMLACWLVRGLAKWRUL PWAYAFGGCVMHSLTLAIHDISINAAFGTGF AARNWLAVFANL-PGGVPYAASFKKYHVDH HRYLGGGLDUDVPTRLEGWFCTPARKLLL WLVLQPFFYSLRPLCVHPKAVTRMEVLNTLV QLA ARNRWLAVFANL-PGGVPYAASFKKYHVDH HRYLGGGLDUDVPTRLEGWFCTPARKLLL WLVLQPFFYSLRPLCVHPKAVTRMEVLNTLV QLA LESYRPDTDLSREDTGCNLQHISDRENDDLN MEFNPSDHPRASTIFLJKSGTDVREKKRSLFIN HHPPGQLARKYSSCSTIFLDDSTVSQPNLKYTI KCVALAYYHIKNRDPDGRMLDDFEDNLHPH SKSEVPPDVDKINFPBGKQNTRFVRTLFSAAQL TAECAIVTLVYLERLLTYAEIDICPANWKRIV LIGAILLASK VWDDQAVWNVPYCQILKDITVE DMNELERQFLELLQFNINVPSSVYAKYYDL RSLAEANNLSPHLEPLSRRFAHKLEAISRICED KYKDLRRSARKRSASADNLTLPRWPSDIS KYRDLRRSARKRSASADNLTLPRWPSDIS KYRDLRRSARKRSASADNLTLPRWPSDIS KYRDLRRSARKRSASADNLTLPRWPSDIS KYRDLRRSARKRSASADNLTLPRWPSDIS KYRDLRRSARKRSASADNLTLPRWPSDIS KYRDLRRSARKRSASADNLTLPRWPSDIS KYRDLRRSARKRSASADNLTLPRWPSDIS KYRDLRRSARKRSASADNLTLPRWPSDIS KYRDLRRSARKRSASADNLTLPRWPSDIS KYRDLRSARKRSASADNLTLPRWPSDIS KYRDLRSARKRSASADNLTLPRWPSDIS KYRDLRSARKRSASADNLTLPRWPSDIS KYRDLRSARKRSASADNLTLPRWPSDIS KYRDLRSARKRSASADNLTLPRWPSDIS KYRDLRSARKRSASADNLTLPRWPSDIS KYRDLRSARKRSASADNLTLPRWPSDIS KYRDLRSARKRSASADNLTLPRWPSDIS KYRDLRSARKRSASADNLTLPRWPSDIS KYRDLRSARKRSASADNLTLPRWPSDIS KYRDLRSARKRSASADNLTLPRWPSDIS KYRDLRSARKRSASADNLTLPRWPSDIS KYRDLRSARKRSASADNLTLPRWGFELLYWKAT KAVGIKMGSLSTANVEFCLDVFRELISKNYA KAVGIKMGSLSTANVEFCLDVFRELSKNYA KAVGIKMGSLSTANVEFCLDVFRELSKNYA KAVGIKMGSLSTANVEFCLDVFRELSKNYA KAVGIKMGSLSTANVEFCLDVFRELSKNYA KAVGIKMGSLSTANVEFCLDVFRELSKNYA KAVGIKMGSLSTANVEFCLDVFRELSKNYA KAVGIKMGSLSTANVEFCLDVFRELSKNYA KAVGIKMGSLSTANVEFCLDVFRELSKNYA KAVGIKMGSLSTANVEFC							GRCAVALLSGVCTLVSTHVCVGSGCPGAAGT
GAGPAPGRAMRSTITLLALLALULLYU-VIGAL VFRALEOPHEQOAGRELIGEVREKERAHPCV SDQELGLLIKEVADALGGGADPETNSTSNSS SAWDLGSAFFFSGTITTTIGGGGDWHYGGK ELPHGGRCRETEGSQVAPRLPASPLCPGYGN VALKTDAGRLFGIFVALVGIPLFGILLAGVGD RLGSSLRHGIGGHEAFILKMPPPLVRVLSA A 9873 2 652 FVVPSPCGGIFGRAPNGASRPTMGNSASRNDE EWYTDQPHTQRKELLAKYPAIKALMRPDP RLKWAVI-VLVLVOMLACVKJLAWFWLI FWAYAFGGCVHISLTLAIHDISINAAFGTGR AARNEWLAVFANLPEGGYPASFKLYRNHH HRYLGGDGLDVDVPTRLEGWFFCTPARKLL WLVLQPFFYSLRPLCVHPKAVTRMEVLNTLV QLA 1251 2601 A 9875 150 1209 PVIMPLHSPGDIVRPSCCVSSSPKLRNAHSR LESYRPTDLSREDTGCNLQHISDRENDDLN MENPSDHPRASTIFLSKSQTDVREKKSLFN HHPPGQLAKXYSCSTHLDDSTVSQPNLKYTI KCVALAIYYHKNRDPDGRMILDIFDENLHPI SKSEVPPDYBKINFPGQKQTVRFVRTLFSAQI TARCAIVTLVYLERLLTYAEDICPANWKRIV LGAILLASKVWDDQAVWNVDYCQLKDITVE DMNELREQFLELLQPNINVPSSVYAKYYPDL RSLAEANNLSPFLEPLSRERAHKLEAISRLCEDKYKDLRSKSTANVEFCLDVFKEILSNNIG DNIFFSSLSLYALSSACLIAGAGETEEQLEKV KNSEVVESPPRSLSCSRSGSAKILISLYQ KEQAELLYGLYCCQCDLTSSHPSSVPAMSSC 1252 2602 A 9880 180 388 KEQAELLYGLYCCQCDLTSSHPSSVPAMSSC 1253 2603 A 9880 180 388 KEQAELLYGLYCCQCDLTSSHPSSVPAMSSC 1254 2604 A 9881 19 494 VISFQIITDIMDSSTAHSPVFLVPPPEITASEYE STELSATTFSTQSPLQKLFARKMKILGTIQILF GIMTFSGVSLFTLLKPYRFFLSGYPFWG SVI_FINSGAFLIAVKKTTETLIILSRIMMFLSA LGAIAGIILLTFEFPRSKHL 1255 2605 A 9896 72 386 RPGREGRDCFQAPPLGIGGRQTDMMHHPLT GATCVGLPNVGMCPQLSGALTFMYLQGINQ EATVAPDTMAQPYASAQPAPQNGIPGETTA PHPHPAPEVTGGTT			1	1			-
VFRALEOPHEQOAQREIGEVREKFLRAHPCV SDQBLGILLIKEWADALGGAPETNISTNSSIS SAWDLGSAFFFSGTIITTIGGGDWHVGGGK ELPHGGRCRETEGSQVAPRLPASPLCPGYGN VAIATDAGRIFCITYALVGUPLFGILLAGVGD RLGSSLRHGIGHIEAIFLKWHVPELVYLSA MLFLLIGGLLPLVTIPTIFVCYMEDWSKLEAITY FVIVILITUTIVGFGDYVA 1250 2600 A 9873 2 652		{	1	1			
SDQELGI_LIKEV_ADA_LGGGADPETNSTNSSISS			1				
SAWDLGSAPFFSGTIITTIGGGGDWHVGGGK ELPHGGRCRETEGSQVAPRLPASPLCPGYGN VALRTDAGRLEGIFYALVGIPLFGILLAGVGD RLGSSLRHGIGHIEAIFLKWHVPPELVRUSA MLFILLIGCLIFVLTPITYPCYMEDWSKLEAIY FVIVTLTTVGFGDYVA FVIVTLTTVGFGDYVA FVIVTLTTVGFGDYVA FVIVTLTTVGFGDYVA FVIVTLTTVGFGDYVA FVIVTLTTVGFGDYVA FVIVTLTTVGFGDYVA FVIVTLTTVGFGDYVA FVIVTLTTVGFGDYVA FVIVTLTTVGFGDYVA FVIVTLTTVGFGDYVA FVIVTLTTVGFGDYVA FVIVTLTTVGFGDYVA FVIVTLTVGFGDYVA FVIVTLTVGFGDYVA FVIVTLTVGFGDYVA FVIVTLTVGFGDYVA FVIVTLAHIDISHNAAFGTGR AARNEWLAVFANLPEGVPYAASFGKYHVDH HRYLGGDGLDVDYTHLGGWFFCTPARKLL WLVLQPFFYSLRPLCVHPKAVTRMEVLNTLV QLA LSYRPDTDLSREDTGCNLQHISDRENDDLN MEFNPSDHPASTITLLSKGJTDVREKKRSLFIN HPPGQLARKYSSCSTIFLDDSTVSQPNLKYTI KCVALAIYYHIKNRDPDGRMLLDIDFENLHPH SKSEVPPDYDKHNPEGKQNIZFVRTLFSAAQL TAECAIVTLVYLERLLTYAEDICPANWKRIV LGAILLASKWDDQAVAWNDYQQILKDITVE DMNELERQFLELLQFNINNPSSVYAKYYPDL RSLAEANNLSFPLERSREAHKLEAISNLCED KYKDLRRSARKRSASADNLTLPR WSPAIIS KRPDSRPPAQYRAGPTRPRTTRGCLLYWKAT LGAILLASK WDDQAVAWNDYQQILKDITVE DMNESSLSLLYALSSWVLLGARGETEGOLEKV WNSSEVCSEPRSLSCSRSGSAKLILSLYQ KRPDSRPPAQYRAGPTRPRTTRGCLLYWKAT KAVGIKMGSLSTANVEGLDYKELNSNNIG DNIFSSLSLLYALSSWVLLGARGETEGOLEKV WNSSEVCSEPRSLSCSRSGSAKLILSLYQ LGAIGILLYGLYQCQCDLTJSSHPSSVPAMSSC NFTHATTVLIGIPGLEKAHFWVGPPLLSMYVA AMFGNC STELSATTFSTQSPLQKLFARKMKLGTIQLT GIMTTSFGVIELTALKTVPRPPFIFLSGYPFWG SVLFINSGAFLIAVKRKTTETLILISRIMNIFLSA LGAIAGIILLTFEHPRSKHLGTIQLT GATCVGLPNVOMCPQLSGALTFMVLQGNQ SATVAPDTMAQPYASAGAPAPQNGIGGEYTA PHPHPAPEYTGQTT GATCVGLPNVOMCPQLSGALTFMVLQGNG SATVAPDTMAQPYASAGAPAPQNGIGGEYTA PHPHPAPEYTGQTT PHPH			1			[
			1				1
VALATDAGRLECIFYALVGIPLEGILAGVGID RLGSSLRHGIGHIEALIFLKWHVPPELVRVLSA MLFLLIGCLLFVLTPIFYPCYMEDWSKLEAIY FVIVILTITVGFGDYVA FVIVILTITVGFGDYVA FVIVILTITVGFGDYVA FVVYFDGGGFGRAPNGASRPIMGNSASRNDF EWYTDQPHTQRRKELLAKYPAIKALMRPPP RLKWAVLVLVQMLACWLVRGLAWRWLI FWAYAFGGCVNISLITLAHIDISHNAAFGTGR AARNRWLAVFANLPEGVFYAASFKYHVDH HRYLGGGGLDVDYFILEGWFFCTPARKLL WLVLQPFFYSLRPLCVHPKAVTRMEVLNTLV QLA							
RLGSSLRHGIGHIEAIFLKWHVPPELVRVLSA MFLLIGCLLFVLTPTFVCYMEDWSKLEAIY FVIVILTITVGFGDYVA 1250 2600 A 9873 2 652 FVVPSPCGGIPGRAPNGASRFTMGNASRNUL FWAYAFGGCWHSLTLAIHUISHNAAFGTGR AARNRWLAVFANLPEGVPYAASFKKYHVDH REWYATDOPHTORREELAKYPAKALMRPDP RLKWAVLVLVQMLACWLVRGLAWRWLL FWAYAFGGCWHSLTLAIHUISHNAAFGTGR AARNRWLAVFANLPEGVPYAASFKKYHVDH REWYLGGDGLDDVDYPTRLEGWFFCTPARKLL WLVVLQPFFYSLRPLCVHPKAVTRMEVLNTLV QLA 1251 2601 A 9875 150 1209 PVIMPLHFSPGDIVRFSCCVSSSPKLRRNAHSR LESYRPDTDLSREDTGCNLQHISDRENIDDLN MEFNPSDHPRASTIFLSKSQTDVREKRKSLFIN HHPPGQIARKYSSCSTIFLDDSTVSQPNLKYTI KCVALAIYYHIKNRDPDGRMLLDIFDENLHPL SKSEVPPDYDKINPEQKQTYRVRTLFSAAQL TAECAIVTLVYLERLLTYAEIDICPANWKRV LGAILLASKVWDDQAVWNVDYGQILKDITVE DNNELERQFLELLQPNINVPSSVYAKYYFDL RSLAEANNLSPPLEPLSRERAHKLEAISRLCED KYKDLRRSARKSSAADNLTLPRWSPAIIS REPUSPRPAQYRAGPPRTRGCELLYWKAT KAVGIKMGSLSTANVEFCLDVFKELNSNNIG DNIFFSSLSLLYALSMVNLGARGETEGQLEXV WNSSEVSESPRSLSCSRGGSAKLILSLYQ WNSSEVSESPRSLSCSRGSGKLILSLYQ WNSSEVSESPRSLSCSRGSGKLILSLYQ WNSSEVSESPRSLSCSRGSGKLILSLYQ VISSQITDTTIMDSSTAHSPVFLVPPEITASEY STELSATTFSTQSPLQSLFARKMKILGTIQILF GIMTFSFQVIFLTLLKPYPPFETASEY GIMTFSFGVIFLTLLKPYPPFETASEY STELSATTFSTQSPLQSLFARKMKILGTIQILF GIMTFSFGVIFLTLLKPYPPFETASEY STELSATTFSTQSPLQSLFARKMKILGTIQILF GIMTFSFGVIFLTLLKPYPPFETASEY STELSATTFSTQSPLQSLFARKMKILGTIQILF GIMTFSFGVIFLTLLKPYPPFETASEY CATAVAPTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTCQTT GATCVGLPNVGMCPJSAATTFMYLQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTQGTT GATCVGLPNVGMCPJSAATTFMYLQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTQGTT GATCVGLPNVGMCPJSAATTFMYLQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTQGTT GATCVGLPNVGMCPJSAATTFMYLQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTQGTT GATCVGLPNVGMCPJSAATTFMYLQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTQGTT PHPHPAPEYTQGTT PHPHPAPEYTQGTT PHPHPAPEYTQGTT PHPHPAPEYTQGTT PHPHPAPEYTGGTT PHPHPAPEYTGGTT PHPHPAPEYTGGTT PHPHPAPEYTGGTT PHPHPAPEYTGGTT PHPHPAPEYTGGTT PHPHP		l	ĺ	ļ			
M_FILIGCLI_PVLTPTTPCYMEDWSKLEAIY		ľ					1
1250 2600 A 9873 2 652 FVVPSPCGGIPGRAPNGASRPTMGNSASRNDF RLKWAVLVLV, VOMLACWLVRGLAWRWL FWYAPGGCVNHSLIALHIJSHNAAFGTGR AARNRWLAVFANLPEGVPYAASFKKYHVDH HRYLGGDGLDVDVPTRLEGWFFCTPARKLL WLVLQFFYSLRPLCHAHDISHNAAFGTGR AARNRWLAVFANLPEGVPYAASFKKYHVDH HRYLGGDGLDVDVPTRLEGWFFCTPARKLL WLVLQFFYSLRPLCHKAVTRMEVLNTLV QLA PVIMPLHFSPGDIVRPSCCVSSSPKLRRNAHSR LESYRPDTDLSREDT GCNLQHISDRENIDDLN MERNPSDHFRASTIFLSKSGTDDVREKKRSLFIN HIPPGGJARKYSSCSTIFLDDSTVSGPNLKYTI KCVALAIYYHIKNRDPDGRMLLDIFDENLHPL SKSEVPDYDVRINPEDGVRYKRVRILFSAAQL TAECAIVTLVYLERLLTYAEIDICPANWKRIV LGAILLASKVWDDQAVWNVDYCQILKDITVE DMNELERGPLELLGYBLTYAEIDICPANWKRIV LGAILLASKVWDDQAVWNVDYCQILKDITVE DMNELERGPLELLGYBLTYVETILFRSAFAIS RSLAEANNLSFPLEPLSRERAHKLEAISRLCED KYKDLRRSARKRSASADNLTLPRWSPAIIS RSLAEANNLSFPLEPLSRERAHKLEAISRLCED KYKDLRRSARKRSASADNLTLPRWSPAIIS RSLAEANNLSFPLEPLSRERAHKLEAISRLCED KYKDLRRSARKRSASADNLTLPRWSPAIIS RSLAEANNLSFPLEPLSRERAHKLEAISRLCED KYKDLRRSARKRSASADNLTLPRWSPAIIS RSPORTPRRTRGCELLYWKAT KAAGIKMGSLSTANVEFCLDVFKELNSINNIG DNIFFSSLSLLYALSMVLLGARGETEEQLEKV WNSSEVCSEPRSLSCSRGSAKLILSLYQ WNSSEVCSEPRSLSCSRGSAKLILSLYQ WNSSEVCSEPRSLSCSRGSAKLILSLYQ WNSSEVCSEPRSLSCSRGSAKLILSLYQ STELSATTFSTQSPLQKLFARKMKILGTIQILF GIMTFSFGVIFLFTLLKPYPPETTASEYE STELSATTFSTQSPLQKLFARKMKILGTIQILF GIMTFSFGVIFLFTLLKPYPRFPFIELSGYPFWG SVLFINSGAFLIAVKRKTTETLIILISRIMNFLSA LGAIAGIILLTTEFHPRSKLHL GIMTFSFGVIFLFTLLKPYPRFPETSGYPFWG SVLFINSGAFLIAVKRKTTETLIILISRIMNFLSA LGAIAGIILLTTEFHPRSKLHL LGAIAGIILLTTEFHPRSKLHL LGAIAGIILLTTEFHPRSKLHL LGAIAGIILLTTEFHPRSKLHL GIMTFSFGVIFLFTLLKPYPRFPETGGYPFWG SVLFINSGAFLIAVKRKTTETLIILISRIMNFLSA LGAIAGIILLTTEFHPRSKLHL LGAITCHPVBMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT GATCVCLPPVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT GATCVCLPPVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT GATCVCLPPVGMCPCLSGALTFMYLQGGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT GATCVCLPPVGMCPCL		ļ	ļ	1		ļ	
1250		ļ	1	İ			
EWYYTDQPHTQRRKEILAKYPAIKALMRPDP RLKWAVLVLVQMLACWLVRGLAWRVLL FWAYAFGGCVMYALTALHIDISHNAAFGTGG AARNRWLAVFANDLPEGVPYAASFKKYHVDH HRYLGGDGLDVDVFTRLEGWFFCTPARKLL WLVLQPFFYSLRFLCVHPKAVTRMEVLNTLV QLA			L				
RIKWAVI-VI-VI-VOMLACWI-VRGI-AWRWILI	1250	2600	Α	9873	2	652	
FWAYARGGCVNHSLTLAHDISHNAAFGTGR AARNRUAVFANIPEGVPYAASFKKYHVDH HRYLGGDGLDVDVPTRLEGWFFCTPARKLL WLVLQPFFYSLRPLCVHPKAVTRMEVLNTLV QLA QLA PVIMPLHFSPGDIVRPSCCVSSSPKLRRNAHSR LESYRPDTDLSREDTGCNLQHISDRENIDDLN MEFNPSDHPRASTIFLSKSQTDVREKRKSLFIN HHPPGQIARKYSSCSTIFLDDSTVSQPNLKYTI KCVALAIYYHIKNRDPDGRMLLDIFDENLHPI SKSEVPPDYDKHNPEQKQIYRFVRTLFSAAQI TAECAIVTLVYLERLITVAEIDICPANWRIV LGAILLASKVWDDQAVWNVDYCQILKDITVE DMNELERQFLELLQFNINVPSSVYAKYYFDL RSLAEANNLSPLEPLSRERAHKLEAISRLCED KYKDLRRSARKRSASADNLTLPRWSPAIIS TAECAIVTLVYLIGRAGFTEEQLEKV WNSSEVCSEPRSLSCSRGSAKLILSLYQ WNSSEVCSEPRSLSCSRGSAKLILSLYQ WNSSEVCSEPRSLSCSRGSGSAKLILSLYQ WNSSEVCSEPRSLSCSRGSGSAKLILSLYQ WNSSEVCSEPRSLSCSRGSGSAKLILSLYQ WISFQITTDTIMDSSTAHSPVFLVFPPEITASEYE STELSATTFSTQSPLQKLFARKMKILGTIQILF GIMTFSFGVIFLFTLLKPYPRFPFIFLSGYPFWG SVLFINSGAFILAVKRKTTETLIILSRIMNFLSA LGALAGIILLTFEHPRSKLHL GATCVGLPNVGMCPQLSGALTFMYLQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAEEYTGQTT P		ı	1	ŀ			
AARNRWLAVFANLEGVPYAASFKKYHVDH		,	1				<u>-</u>
HRYLGGDGLDVDVPTRLEGWFFCTPARKLL WLVLQPFFYSLRPLCVHPKAVTRMEVLNTLV QLA 1251 2601 A 9875 150 1209 PVIMPLHFSPGDIVRPSCCVSSSPKLRRNAHSR LESYRPDTDLSREDTGCNLQHISDRENIDDLN MEFNPSDHPRASTIFLSKSQTDVREKKKSLFIN HHPPGQIARKYSSCSTIFLDDSTVSQPNLKYTI KCVALAIYYHIKNRDPOGRMLLDIFDENLHPL SKSEVPPDYDKHNFEQKQIYRFVRTLFSAAQL TAECAIVTLVYLERLLTYAEIDICPANWKRIV LGAILLASK VWDDQAVWNVDYCQILKDITVE DMNELERQFLELLQFNINVPSSVYAKYYFDL RSLAEANNLSFPLEPLSRERAHKLEAISRLCED KYKDLRSAKKRSASADNLTLPRWSPAIIS 1252 2602 A 9879 6 376 KRPDSRPPAYRAGPTPRTRGCELLYWKAT KAVGIKMGSLSTANVEFCLDVFKELNSNNIG DNIFFSSLSLLYALSMVLLGARGETEEQLEKV WNSSEVCSPRSLSCSRSGSAKLILSLYQ WNSSEVCSPRSLSCSRSGSAKLILSLYQ 1253 2603 A 9880 180 388 KEQAELLYGLYCQCDLTLSSHPSSVPAMSSC NFTHATFVLIGIPGLEKAHFWVGFPLLSMYVA AMFGNC AMFGNC STELSATTFSTQSPLQKLFAKKMLGTIQILF GIMTFSFGVIFLFTLLKFPYRFPFIFLSGYPFWG SVLFINSGAFLIAVKRKTTETLIILSRIMNFLSA LGAIAGIILLTFEFHPRSKLHL LGAIAGIILLTFEFHPRSKLHL GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT SGGPAGLLHRPVLPKMGLSGLLPILVFFILLG DIQEPGHAEGIILGKPCPKIKVECEVEEDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE PRGCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE PVINCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE PVINCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE PVINCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE PVINCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE PVINCPENMKCCPFSRGKKCLDFRKVSLTLYH EELE PVINCPENMKCCPFSRGKKCLDFRKVSLTLYH EELE PVINCPENMKCCPFSRGKKCLDFRKVSLTLYH EELE PVINCPENMKCCPFSRGKKCLDFRKVSLTLYH EELE PVINCPENMKCCPFSRGKKCLDFRKVSLTLYH EELE PVINCPENMKCCPFSRGKKCLDFRKVSLTLYH EELE PVINCPENMKCCPFSRGKKCLDFRKVSLTLYH EELE PVINCPENMKCCPFSRGKKCLDFRKVSLTLYH EELE PVINCPENMKCCPFSRGKKCLDFRKVSLTLYH EELE PVINCPENMKCCPFSRGKCLDFRKVSLTLYH EELE PVINCPENMKCCPFSRGKKCLDFRKVSLTLYH EELE PVINCPENMKCCPFSRGKCLDFRKVSLTLYH EELE PVINCPENMKCCPFSRGKCLDFRKVSLTLYH EELE PVINCPENMKCCPFSRGKCLDFRKVSLTLYH EELE PVINCPENMKCCPFSRGKCLDFRKVSLTLYH EELE PVINCPENMCCPPTSRGKCLDFRKVS		ļ	İ				FWAYAFGGCVNHSLTLAIHDISHNAAFGTGR
WLVLQPFFYSLRPLCVHPKAVTRMEVLNTLV QLA		1		1.		İ	AARNRWLAVFANLPEGVPYAASFKKYHVDH
1251 2601 A 9875 150 1209 PVIMPLHFSPGDIVRPSCCVSSSPKLRRNAHSR LESYRPDTDLSREDTGCNLQHISDRENIDDLN MEFNPSDHPRASTIFLSKSQTDVREKRKSLFIN HHPPGQIARKYSSCSTIFLDDSTVSQPNLKYTI KCVALAIYYHIKNRDPDGRMLLDIFDENLHPI SKSEVPPDYDKHNPEQKQIYRFVRTLFSAAQL TAECAIVTLVYLERLLTYAEIDICPANWKRIV LGAILLASK VWDDQAVWNVDYCQILKDITVE DMNELERQFLELLQFNINVPSSVYAKYYFDL RSLAEANNLSFPLEPLSRERAHKLEAISRLCED KYKDLRSARKRSASADNLTLPR WSPAIIS TAPECAIVTLYLASHANLSFPLEPLSRERAHKLEAISRLCED KYKDLRSARKRSASADNLTLPR WSPAIIS TAPECAIVTLYLASHANLSFPLEPLSRERAHKLEAISRLCED KYKDLRSARKRSASADNLTLPR WSPAIIS TAPECAIVTLYLASHANLSFPLEPLSRERAHKLEAISRLCED KYKDLRSARKRSASADNLTLPR WSPAIIS TAPECAIVTLYLASHANLSFPLEPLSREAHKLEAISRLCED KYKDLRSARKRSASADNLTLPR WSPAIIS TAPECAIVTLYLASHANLSFPLEPLSREAHKLEAISRLCED KYKDLRSARKRSASADNLTLPR WSPAIIS TAPECAIVTLYLASHANLSFPLEPLSREAHKLEAISRLCED KYKDLRSARKRSASADNLTLPR WSPAIIS TAPECAIVTLYLASHANLSFPLEPLSREAHKLEAISRLCED KYKDLRSARKRSASADNLTLPR WSPAIIS TAPECAIVTLYLASHANLSFPLEPLSREAHKLEAISRLCED KYKDLRSARKRSASADNLTLPR WSPAIIS TAPECAIVTLYLASHANLSFPLEPLSSANTHSCLLSYN		1	ļ	1			HRYLGGDGLDVDVPTRLEGWFFCTPARKLL
1251							WLVLQPFFYSLRPLCVHPKAVTRMEVLNTLV
LESYRPDTDLSREDTGCNLQHISDRENIDDLN MEFNPSDHPRASTIFLSKSQTDVREKRKSLFIN HHPPGQIARKYSSCSTIFLDDSTVSQPNLKYTI KCVALAITYHIKNRDPDGRMLDIFDENLHPI SKSEVPPDYDKHNPEQKQIVRFVRTLFSAAQI TAECAIVTLVYLERLLTYAEIDICPANWKRIV LGAILLASK WWDDQA VWNVDYCQILKDITVE DMNELERQFLELLQFNINVPSSVYAKYYFDL RSLAEANNLSFPLEPLSRERAHKLEAISRLCED KYYDLRSARKRSASADNLTLPRWSPAIIS 1252 2602 A 9879 6 376 KRPDSRPPAQVRAGPTRPRTRGCELLYWKAT KAVGIKMGSLSTANVEFCLDVFKELNSNNIG DNIFFSSLSLLYALSMVLLGARGETEEQLEKV WNSSEVCSEPRSLSCSRSGSKLILSLYQ 1253 2603 A 9880 180 388 KEQAELLYGLYCQCDLTLSSHPSSVPAMSSC NFTHATFVLIGIPGLEKAHFWVGFPLLSMYVA AMFGNC 1254 2604 A 9881 19 494 VISSQITDTIMDSSTAHSPVFLVFPPEITASEYE STELSATTFSTQSPLQKLFARKMKILGTIQILF GIMTSFGVFFLFTLLKPYPRFFPIFLSGYPFWG SVLFINSGAFLIAVKRKTTETLIILSRIMNFLSA LGAIAGIILTFEFHPRSKLHL 1255 2605 A 9896 72 386 RPGREQRDCFQAPPLGLGGRQTDMMHHPLT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT 1256 2606 A 9902 95 399 SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHAEGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE			1				QLA
MEFNPSDHPRASTIFLSKSQTDVREKRKSLFIN HHPPGQIARKYSSCSTIFLDDSTVSQPNLKYTI KCVALAIYYHIKNRDPDGRMLLDIFDENLHPI SKSEVPPDYDKHNPEQKQIYRFVRTLFSAAQL TAECAIVTLVYLERLLTYAEIDICPANWKRIV LGAILLASKVWDDQAVWNDYCQILKDITVE DMNELERQFLELLQFNINVPSSVYAKYYFDL RSLAEANNLSFPLEPLSRERAHKLEAISRLCED KYKDLRRSARKRSASADNLTLPRWSPAIIS KYKDLRRSARKRSASADNLTLPRWSPAIIS KAPDSRPPAQYRAGPTRPRTRGCELLYWKAT KAVGIKMGSLSTANVEFCLDVFKELNSNNIG DNIFFSSLSLLYALSMVLLGARGETEEQLEKV WNSSEVCSEPRSLSCSRSGSAKLILSLYQ 1253 2603 A 9880 180 388 KEQAELLYGLYCQCDLTLSSHPSSVPAMSSC NFTHATFVLIGIPGLEKAHFWVGFPLLSMYVA AMFGONC 1254 2604 A 9881 19 494 VISFQIITDTIMDSSTAHSPVFLVFPPEITASEYE STELSATTFSTQSPLQKLFARKMKILGTIQILF GIMTFSFGVIFLFTLLKPYPRFFFILSGYPFWG SVLFINSGAFLIAVKRKTTETLIILSRIMNFLSA LGAIAGIILLTFEFHPRSKLHL 1255 2605 A 9896 72 386 RPGREQRDCFQAPPLGLGGRQTDMMHPLTI GATCVGLPNVGMCPQLSSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT 1256 2606 A 9902 95 399 SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHAEGIIGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE	1251	2601	Α	9875	150	1209	PVIMPLHFSPGDIVRPSCCVSSSPKLRRNAHSR
MEFNPSDHPRASTIFLSKSQTDVREKRKSLFIN HHPPGQIARKYSSCSTIFLDDSTVSQPNLKYTI KCVALAIYYHIKNRDPDGRMLLDIFDENLHPI SKSEVPPDYDKHNPEQKQIYRFVRTLFSAAQL TAECAIVTLVYLERLLTYAEIDICPANWKRIV LGAILLASKVWDDQAVWNDYCQILKDITVE DMNELERQFLELLQFNINVPSSVYAKYYFDL RSLAEANNLSFPLEPLSRERAHKLEAISRLCED KYKDLRRSARKRSASADNLTLPRWSPAIIS KYKDLRRSARKRSASADNLTLPRWSPAIIS KAPDSRPPAQYRAGPTRPRTRGCELLYWKAT KAVGIKMGSLSTANVEFCLDVFKELNSNNIG DNIFFSSLSLLYALSMVLLGARGETEEQLEKV WNSSEVCSEPRSLSCSRSGSAKLILSLYQ 1253 2603 A 9880 180 388 KEQAELLYGLYCQCDLTLSSHPSSVPAMSSC NFTHATFVLIGIPGLEKAHFWVGFPLLSMYVA AMFGONC 1254 2604 A 9881 19 494 VISFQIITDTIMDSSTAHSPVFLVFPPEITASEYE STELSATTFSTQSPLQKLFARKMKILGTIQILF GIMTFSFGVIFLFTLLKPYPRFFFILSGYPFWG SVLFINSGAFLIAVKRKTTETLIILSRIMNFLSA LGAIAGIILLTFEFHPRSKLHL 1255 2605 A 9896 72 386 RPGREQRDCFQAPPLGLGGRQTDMMHPLTI GATCVGLPNVGMCPQLSSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT 1256 2606 A 9902 95 399 SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHAEGIIGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE		Ī	j				
HHPPGQIARKYSSCSTIFLDDSTVSQPNLKYTI KCVALAIYYHIKNRDPDGRMLLDIFDENLHPI SKSEVPPDYDKHNPEQKQIYRFVRTLFSAAQI TAECAIVTLVYLERLITYAEIDICPANWKRIV LGAILLASK VWDDQAVWNVDYCQILKDITVE DMNELERQFLELQFNINVPSSVYAKYYFDL RSLAEANNLSFPLEPLSREAHKLEAISRLCED KYKDLRRSARKRSASADNLTLPRWSPAIIS 1252 2602 A 9879 6 376 KRPDSRPPAQYRAGPTRPRTRGCELLYWKAT KAVGIKMGSLSTANVEFCLDVFKELNSNNIG DNIFFSSLSLLYALSMVLLGARGETEGQLEKV WNSSEVCSEPRSLSCSRSGSAKLILSLYQ 1253 2603 A 9880 180 388 KEQAELLYGLYQCODLTLSSHPSSVPAMSSC NFTHATFVLIGIPGLEKAHFWVGFPLLSMYVA AMFGNC 1254 2604 A 9881 19 494 VISFQIITDTIMDSSTAHSPVFLVFPPEITASEYE STELSATTFSTQSPLQKLFARKMKILGTIQILF GIMTFSFGVIFLFTLLKPYPFFFFIFLSGYFFWG SVLFINSGAFLIAVKRKTTETLIILSRIMNFLSA LGAIAGIILLTFEFHPRSKLHL 1255 2605 A 9896 72 386 RPGREQRDCFQAPPLGLGGRQTDMMHHPLT GATCVGLPNVGMCPQLSGALTFMYLQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT 1256 2606 A 9902 95 399 SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHAEGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE							
KCVALAIYYHIKNRDPDGRMLLDIFDENLHPL SKSEVPPDYDKHNPEQKQIYRFVRTLFSAAQL TAECAIVTLYYLERLLTYAEIDICPANWKRIV LGAILLASKVWDDQAVWNVDYCQILKDITVE DMNELERQFLELLQFNINVPSSVYAKYYFDL RSLAEANNLSFPLEPLSRERAHKLEAISR.CED KYKDLRRSARKRSASADNLTLPRWSPAIIS 1252 2602 A 9879 6 376 KRPDSRPPAQYRAGPTRPRTRGCELLYWKAT KAVGIKMGSLSTANVEFCLDVFKELNSNNIG DNIFFSSLSLLYALSMVLLGARGETEEQLEKV WNSSEVCSEPRSLSCSRGSGAKLILSLYQ WNSSEVCSEPRSLSCSRGSGAKLILSLYQ WNSSEVCSEPRSLSCSRGSGAKLILSLYQ WNSSEVCSEPRSLSCSRGSGAKLILSLYQ WNSSEVCSEPRSLSCSRGSGAKLILSLYQ WNSSEVCSEPRSLSCSRGSGAKLILSLYQ WNSSEVCSEPRSLSCSRGSGAKLILSLYQ WNSSEVCSEPRSLSCSRGSGAKLILSLYQ WNSSEVCSEPRSLSCSRGSGAKLILSLYQ WNSSEVCSEPRSLSCSRGSAKLILSLYQ WNSSEVCSEPRSLSCSRGSAKLILSLYQ WNSSEVCSEPRSLSCSRGSAKLILSLYQ WNSSEVCSEPRSLSCSRGSAKLILSLYQ WNSSEVCSEPRSLSCSRGSAKLILSLYQ WNSSEVCSEPRSLSCSRGSAKLILSLYQ WNSSEVCSEPRSLSCSRGSAKLILSLYQ WNSSEVCSEPRSLSCSRGSAKLILSLYQ WNSSEVCSEPRSLSCSRGSAKLILSLYQ WNSSEVCSEPRSLSCSRGSAKLILSLYQ WNSSEVCSEPRSLSCSRGSAKLILSLYQ WNSSEVCSEPRSLSCSRGSAKLILSLYQ WNSSEVCSEPRSLSCSRGSAKLILSLYQ WNSSEVCSEPRSLSCSRGSAKLILSLYQ WNSSEVCSEPRSLSCSRGSAKLILSLYQ WNSSEVCSEPRSLSCSRGSAKLILSLYQ WNSSEVCSEPRSLSCSRGSAKLILSLYQ WNSSEVCSEPRSLSCSRGSAKLILSLYQ WNSSEVCSEPSTAGSAKLILSLYQ WNSSEVCSEPSTAGSAKLILSLYQ WNSSEVCSEPSTAGSAKLILSLYQ WNSSEVCSEPSTAGSAKLILSLYQ WNSSEVCSEPSTAGSAKLILSLYQ WNSSEVCSEPSTAGSAKLILSLYQ WNSSEVCSEPSTAGSAKLILSLYQ WNSSEVCSEPSTAGAKCLDFRKVSLTLYH KEELE WSTATATATATATATATATATATATATATATATATATATA		Ì					,
SKSEVPPDYDKHNPEQKQIYRFVRTLFSAAQL TAECAIVTLVYLERLLTYAEIDICPANWKRIV LGAILLASKVWDDQAVWNVDYCQILKDITVE DMNELERQFLELLQFNINVPSSVYAKYYFDL RSLAEANNLSFPLEPLSRERAHKLEAISRLCED KYKDLRRSARKRSASADNLTLPRWSPAIIS 1252 2602 A 9879 6 376 KRPDSRPPAQYRAGPTRPRTRGCELLYWKAT KAVGIKMGSLSTANVEFCLDVFKELNSNNIG DNIFFSSLSLLYALSMVLLGARGETEGQLEKV WNSSEVCSEPRSLSCSRSGSAKLILSLYQ 1253 2603 A 9880 180 388 KEQAELLYGLYCQCDLTLSSHPSSVPAMSSC NFTHATFVLIGIPGLEKAHFWVGFPLLSMYVA AMFGNC 1254 2604 A 9881 19 494 VISFQIITDTIMDSSTAHSPVFLVFPPEITASEYE STELSATTFSTQSPLQKLFARKMKILGTIQILF GIMTFSFGVIFLFTLLKPYPRFFIFLSGYPFWG SVLFINSGAFLIAVKRKTTETLIILSRIMNFLSA LGAIAGIILTTEFHPRSKLHL 1255 2605 A 9896 72 386 RPGREQRDCFQAPPLGGRQTDMMHHPLT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT 1256 2606 A 9902 95 399 SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHAEGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE			İ	i			, -
TAECAIVTLVYLERLLTYAEIDICPANWKRIV LGAILLASKVWDDQAVWNVDYCQILKDITVE DMNELERQFLELLQFNINVPSSVYAKYYFDL RSLAEANNLSFPLEPLSRERAHKLEAISRLCED KYKDLRRSARKRSASADNLTLPRWSPAIIS 1252 2602 A 9879 6 376 KRPDSRPPAQYRAGPTRPRTRGCELLYWKAT KAVGIKMGSLSTANVEFCLDVFKELNSNNIG DNIFFSSLSLLYALSMVLLGARGETEEQLEKV WNSSEVCSEPRSLSCSRSGSAKLILSLYQ 1253 2603 A 9880 180 388 KEQAELLYGLYCQCDLTLSSHPSSVPAMSSC NFTHATFVLIGIPGLEKAHFWVGFPLLSMYVA AMFGNC 1254 2604 A 9881 19 494 VISFQIITDTIMDSSTAHSPVFLVFPPEITASEYE STELSATTFSTQSPLQKLFARKMKILGTIQILF GIMTFSFGVIFLFTLLKPYPRFFIFLSGYPFWG SVLFINSGAFLIAVKRKTTETLIILSRIMNFLSA LGAIAGIILLTFEFHPRSKLHL 1255 2605 A 9896 72 386 RPGREQRDCFQAPPLGGGRQTDMMHHPLT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT 1256 2606 A 9902 95 399 SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHAEGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE							
LGAILLASK VWDDQA VWNVDYCQILKDITVE DMNELERQFLELL QFNINVPSSVYAKYYFDL RSLAEANNLSFPLEPL SRERAHKLEAISRLCED KYKDLRRSARKRSASADNLTLPR WSPAIIS 1252 2602 A 9879 6 S76 KRPDSRPPAQYRAGPTRPRTRGCELLYWKAT KAVGIKMGSLSTANVEFCLDVFKELNSNNIG DNIFFSSLSLLYALSMVLLGARGETEEQLEKV WNSSEVCSEPRSLSCSRSGSAKLILSLYQ 1253 2603 A 9880 180 388 KEQAELLYGLYCQCDLTLSSHPSSVPAMSSC NFTHATFVLIGIPGLEKAHFWVGFPLLSMYVA AMFGNC 1254 2604 A 9881 19 494 VISFQIITDTIMDSSTAHSPVFLVFPPEITASEYE STELSATTFSTQSPLQKLFARKMKILGTIQILF GIMTFSFGVIFLFTLLKPYPRFFFIFLSGYPFWG SVLFINSGAFLIAVKRKTITETLIILSRIMNFLSA LGAIAGIILLTFEFHPRSKLHL 1255 2605 A 9896 72 386 RPGREQRDCFQAPPLGLGGRQTDMMHHPLT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT 1256 2606 A 9902 95 399 SGGPAGLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHAEGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE							
DMNELERQFLELLQFNINVPSSVYAKYYFDL RSLAEANNLSFPLEPLSRERAHKLEAISRLCED KYKDLRRSARKRSASADNLTLPRWSPAIIS 1252 2602 A 9879 6 376 KRPDSRPPAQYRAGPTRPRTRGCELLYWKAT KAVGIKMGSLSTANVEFCLDVFKELNSNNIG DNIFFSSLSLLYALSMVLLGARGETEEQLEKV WNSSEVCSEPRSLSCSRSGSAKLILSLYQ 1253 2603 A 9880 180 388 KEQAELLYGLYCQCDLTLSSHPSSVPAMSSC NFTHATFVLIGIPGLEKAHFWVGFPLLSMYVA AMFGNC 1254 2604 A 9881 19 494 VISFQIITDTIMDSSTAHSPVFLVFPPEITASEYE STELSATTFSTQSPLQKLFARKMKILGTIQILF GIMTFSFGVIFLFTLLKPYPRFPFIFLSGYPFWG SVLFINSGAFLIAVKRKTIETLILSRIMNFLSA LGAIAGIILLTFEFHPRSKLHL 1255 2605 A 9896 72 386 RPGREQRDCFQAPPLGLGGRQTDMMHHPLT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT 1256 2606 A 9902 95 399 SGGPAGLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHAEGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE				1			
RSLAEANNLSFPLEPLSRERAHKLEAISRLCED KYKDLRRSARKRSASADNLTLPRWSPAIIS 1252 2602 A 9879 6 376 KRPDSRPPAQYRAGPTRPRTRGCELLYWKAT KAVGIKMGSLSTANVEFCLDVFKELNSNNIG DNIFFSSLSLLYALSMVLLGARGETEEQLEKV WNSSEVCSEPRSLSCSRSGSAKLILSLYQ 1253 2603 A 9880 180 388 KEQAELLYGLYCQCDLTLSSHPSSVPAMSSC NFTHATFVLIGIPGLEKAHFWVGFPLLSMYVA AMFGNC 1254 2604 A 9881 19 494 VISFQITIDTIMDSSTAHSPVFLVFPPEITASEYE STELSATTFSTQSPLQKLFARKMKILGTIQILF GIMTFSFGVIFLFTLLKPYPRFPFIFLSGYPFWG SVLFINSGAFLIAVKRKTTETLIILSRIMNFLSA LGAIAGIILLTFEFHPRSKLHL 1255 2605 A 9896 72 386 RPGREQRDCFQAPPLGLGGRQTDMMHHPLT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT 1256 2606 A 9902 95 399 SGGPAGLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHAEGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE				1			
KYKDLRRSARKRSASADNLTLPRWSPAIIS			i				l
1252 2602 A 9879 6 376 KRPDSRPPAQYRAGPTRPRTRGCELLYWKAT KAVGIKMGSLSTANVEFCLDVFKELNSNNIG DNIFFSSLSLLYALSMVLLGARGETEEQLEKV WNSSEVCSEPRSLSCSRSGSAKLILSLYQ 1253 2603 A 9880 180 388 KEQAELLYGLYCQCDLTLSSHPSSVPAMSSC NFTHATFVLIGIPGLEKAHFWVGFPLLSMYVA AMFGNC 1254 2604 A 9881 19 494 VISFQIITDTIMDSSTAHSPVFLVFPPEITASEYE STELSATTFSTQSPLQKLFARKMKILGTIQILF GIMTFSFGVIFLTLLKPYPRFPFIFLSGYPFWG SVLFINSGAFLIAVKRKTTETLIILSRIMNFLSA LGAIAGIILLTFEFHPRSKLHL 1255 2605 A 9896 72 386 RPGREQRDCFQAPPLGLGGRQTDMMHHPLT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT 1256 2606 A 9902 95 399 SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHABGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE			1			·	
KAVGIKMGSLSTANVEFCLDVFKELNSNNIG DNIFFSSLSLLYALSMVLLGARGETEEQLEKV WNSSEVCSEPRSLSCSRSGSAKLILSLYQ 1253 2603 A 9880 180 388 KEQAELLYGLYCQCDLTLSSHPSSVPAMSSC NFTHATFVLIGIPGLEKAHFWVGFPLLSMYVA AMFGNC 1254 2604 A 9881 19 494 VISFQIITDTIMDSSTAHSPVFLVFPPEITASEYE STELSATTFSTQSPLQKLFARKMKILGTIQILF GIMTFSFGVIFLTLLKPYPRFPFIFLSGYPFWG SVLFINSGAFLIAVKRKTTETLIILSRIMNFLSA LGAIAGIILLTFEFHPRSKLHL 1255 2605 A 9896 72 386 RPGREQRDCFQAPPLGLGGRQTDMMHHPLT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT 1256 2606 A 9902 95 399 SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHABGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE	1252	2602	 	0070	-	276	
DNIFFSSLSLLYALSMVLLGARGETEEQLEKV WNSSEVCSEPRSLSCSRSGSAKLILSLYQ 1253 2603 A 9880 180 388 KEQAELLYGLYCQCDLTLSSHPSSVPAMSSC NFTHATFVLIGIPGLEKAHFWVGFPLLSMYVA AMFGNC 1254 2604 A 9881 19 494 VISFQIITDTIMDSSTAHSPVFLVFPPEITASEYE STELSATTFSTQSPLQKLFARKMKILGTIQILF GIMTFSFGVIFLFTLLKPYPRFFFIFLSGYPFWG SVLFINSGAFLIAVKRKTTETLIILSRIMNFLSA LGAIAGIILLTFEFHPRSKLHL 1255 2605 A 9896 72 386 RPGREQRDCFQAPPLGLGGRQTDMMHHPLT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT 1256 2606 A 9902 95 399 SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHARGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE	1232	2002	A	7017	0	370	
WNSSEVCSEPRSLSCSRSGSAKLILSLYQ			1				
1253 2603 A 9880 180 388 KEQAELLYGLYCQCDLTLSSHPSSVPAMSSC NFTHATFVLIGIPGLEKAHFWVGFPLLSMYVA AMFGNC 1254 2604 A 9881 19 494 VISFQIITDTIMDSSTAHSPVFLVFPPEITASEYE STELSATTFSTQSPLQKLFARKMKILGTIQILF GIMTFSFGVIFLFTLLKPYPRFPFIFLSGYPFWG SVLFINSGAFLIAVKRKTTETLIILSRIMNFLSA LGAIAGIILLTFEFHPRSKLHL 1255 2605 A 9896 72 386 RPGREQRDCFQAPPLGLGGRQTDMMHHPLT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT 1256 2606 A 9902 95 399 SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHAEGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE			l				l •
NFTHATFVLIGIPGLEKAHFWVGFPLLSMYVA AMFGNC 1254 2604 A 9881 19 494 VISFQIITDTIMDSSTAHSPVFLVFPPEITASEYE STELSATTFSTQSPLQKLFARKMKILGTIQILF GIMTFSFGVIFLFTLLKPYPRFPFIFLSGYPFWG SVLFINSGAFLIAVKRKTTETLIILSRIMNFLSA LGAIAGIILLTFEFHPRSKLHL 1255 2605 A 9896 72 386 RPGREQRDCFQAPPLGLGGRQTDMMHHPLT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT 1256 2606 A 9902 95 399 SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHAEGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE	1000		<u> </u>	L			
AMFGNC 1254 2604 A 9881 19 494 VISFQIITDTIMDSSTAHSPVFLVFPPEITASEYE STELSATTFSTQSPLQKLFARKMKILGTIQILF GIMTFSFGVIFLFTLLKPYPRFPFIFLSGYPFWG SVLFINSGAFLIAVKRKTTETLIILSRIMNFLSA LGAIAGIILLTFEFHPRSKLHL 1255 2605 A 9896 72 386 RPGREQRDCFQAPPLGLGGRQTDMMHHPLT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT 1256 2606 A 9902 95 399 SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHAEGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE	1253	2603	A	9880	180	388	
1254 2604 A 9881 19 494 VISFQIITDTIMDSSTAHSPVFLVFPPEITASEYE STELSATTFSTQSPLQKLFARKMKILGTIQILF GIMTFSFGVIFLFTLLKPYPRFPFIFLSGYPFWG SVLFINSGAFLIAVKRKTTETLIILSRIMNFLSA LGAIAGIILLTFEFHPRSKLHL 1255 2605 A 9896 72 386 RPGREQRDCFQAPPLGLGGRQTDMMHHPLT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT 1256 2606 A 9902 95 399 SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHAEGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE			1				
STELSATTFSTQSPLQKLFARKMKILGTIQILF GIMTFSFGVIFLFTLLKPYPRFPFIFLSGYPFWG SVLFINSGAFLIAVKRKTTETLIILSRIMNFLSA LGAIAGIILLTFEFHPRSKLHL 1255 2605 A 9896 72 386 RPGREQRDCFQAPPLGLGGRQTDMMHHPLT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT 1256 2606 A 9902 95 399 SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHABGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE				<u> </u>	_		AMFGNC
GIMTFSFGVIFLTLLKPYPRFPFIFLSGYPFWG SVLFINSGAFLIAVKRKTTETLIILSRIMNFLSA LGAIAGIILLTFEFHPRSKLHL 1255 2605 A 9896 72 386 RPGREQRDCFQAPPLGLGGRQTDMMHHPLT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT 1256 2606 A 9902 95 399 SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHABGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE	1254	2604	Α	9881	19	494	VISFQIITDTIMDSSTAHSPVFLVFPPEITASEYE
GIMTFSFGVIFLTLLKPYPRFPFIFLSGYPFWG SVLFINSGAFLIAVKRKTTETLIILSRIMNFLSA LGAIAGIILLTFEFHPRSKLHL 1255 2605 A 9896 72 386 RPGREQRDCFQAPPLGLGGRQTDMMHHPLT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT 1256 2606 A 9902 95 399 SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHABGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE			İ				STELSATTFSTQSPLQKLFARKMKILGTIQILF
SVLFINSGAFLIAVKRKTTETLIILSRIMNFLSA LGAIAGIILLTFEFHPRSKLHL 1255 2605 A 9896 72 386 RPGREQRDCFQAPPLGLGGRQTDMMHHPLT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT 1256 2606 A 9902 95 399 SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHAEGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE							GIMTFSFGVIFLFTLLKPYPRFPFIFLSGYPFWG
LGAIAGIILLTFEFHPRSKLHL 1255 2605 A 9896 72 386 RPGREQRDCFQAPPLGLGGRQTDMMHHPLT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT 1256 2606 A 9902 95 399 SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHAEGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE		[1			1
1255 2605 A 9896 72 386 RPGREQRDCFQAPPLGLGGRQTDMMHHPLT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT 1256 2606 A 9902 95 399 SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHAEGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE		1	1				
GATCVĞLPNVĞMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPĞEYTA PHPHPAPEYTGQTT 1256 2606 A 9902 95 399 SĞGPAĞLLHRPVLPKMĞLSĞLLPILVPFILLĞ DİQEPĞHAEĞILĞKPCPKİKVECEVEEİDQCTK PRDCPENMKCCPFSRĞKKCLDFRKVSLTLYH KEELE	1255	2605	A	9896	72	386	
EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT 1256 2606 A 9902 95 399 SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHAEGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE			١.,	1	·-	.	
PHPHPAPEYTGQTT 1256 2606 A 9902 95 399 SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHAEGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE		}	ŀ				
1256 2606 A 9902 95 399 SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHAEGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE				1 1		ľ	, , ,
DIQEPGHAEGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE	1256	2606	A	0002	05	200	
PRĎCPENMKCCPFSRGKKCLDFRKVSLTĽ YH KEELE	1430	2000	А	9902	כע	299	
KEELE		'			İ	l	, ,
						İ	
1257 2607 A 9905 374 459 EHLKSTPNRLGVVAHTCNPSTLGGRGGW	10.5						
	1257	2607	<u>A</u>	9905	374	459	EHLKSTPNRLGVVAHTCNPSTLGGRGGW

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence.	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, I=possible nucleotide deletion, \=possible nucleotide insertion
1258	2608	A	9911	364	1974	AGPGVPAVGGRWASGPGLGGRTLCSGPPDH QRRGPSCGASGDPQCVGSPHPQRARPLLARP GARLLPGHLPSPRPPRLPTGQPPAAAFRGPVR PQGGGHIHPLPTPGGRPCFAVSEGSGSALLLS YLGECGSSSYVTGAACISPVLRCREWFEAGLP WPYERGFLLHQKIALSRYATALEDTVDTSRL FRSRSLREFEEALFCHTKSFPISWDAYWDRND PLRDVDEAAVPVLCICSADDPVCGPPDHTLTT ELFHSNPYFFLLLSRHGGHCGFLRQEPLPAWS HEVILESFRALTEFFRTEERIKGLSRHRASFLG GRRRGGALQRREVSSSSNLEEIFNWKRSYTRL MAAAAGAAAAPGSREPQDRPECGAGHPGPR YYRHPERWLLRPEAFLGPLRTRAPSAEDSQR ERPAARSGPEMRVRYPVVAAVLAPYLALSQD PMVKSSASGQGASGSYNHVREEMLIKAGGA MSRRVVRQSKFRHVFGQAAKADQAYEDIRV SKVTWDSSFCAVNPKFLAIIVEAGGGGAFIVL PLAK
1259	2609	A	9919	693	935	GCFKFIGESTCCWIFPSSVITQCVVAKAPRAA TLSKAERLRSQPGPEQGGSSYRPRTPTAAAIL
1260	2610	A	9921	455	1082	PPRPGRSHRKRKLVSTK QRSCLCSAIEKDGGDVKALYRRSQALEKLGR LDQAVLDLQRCVSLEPKNKVFQEALRNIGGQ IQEKVRYMSSTDAKVEQMFQILLDPEEKGTE KKQKASQNLVVLAREDAGAEKIFRSNGVQLL QRLLDMGETDLMLAALRTLVGICSEHQSRTV ATLSILGTRRVVSILGVESQAVSLAACHLLOV
1261	2611	A	9928	1	438	MFDALKEGVKKGFRGKEGAIIV GFRGAEAPGAAQAPKKKKPRPTEGGPGAGSG RGKDPYRGPTLLHQPKPPKDEFLSSLESYEIAF
						PTRVDHNGALLAFSPPPPORORRGTGATAES RLFYKEASPSTHFLLNLTRSSRLLAGHVSVEY WTREGLAWQRADRPHCLYA
1262	2612	A	9931	168	435	AAEMGRAGAAAVIPGLALLWAVGLGGPPPA PPRLPFCLQELQGRHALHTFSLERTCSYQDFL WADEGRLLHVGAQDLATWHTLSPLGLW
1263	2613	A	9938	247	488	RMSATSVDQRPKGQGNKVSVQNGSIHQKDG CNDDDFEPYLRSPDNQSNSYPPMSDPYMPGY YAPSIGFPYSLGEAAWSQL
1264	2614	A	9941	61	277	ESIGLTALGPRRRPWEHRWSDPITLKMKGWG WLALLLGALLGTAWARRSQDLHCGACKAVR
1265	2615	A	9956	2	522 · ;	RRVRQFNIYDY FVASEVSKMPVPASWPHPPGPFLLTLLGLT EVAGEEELQMIQPEKLLLVTVGKTATLHCTV TSLLPVGPVLWFRGVGPGRELIYNQKEGHFP RVTTVSDLTKRNNMDFSIRISSITPADVGTYY CVKFRKGSPDHVEFKSGAGTELSVRGEYSVG FLSQVWWWLSSHPFMN
1266	2616	A	10002	243	387	PKNNACHLLFTAVCQPRCKHGECIGPNKCKC HPGYAGKTCNQGRKTV
1267	2617	A	10004	36	707	LPAPASTWSVARETMASSSVPPATVSAATAG PGPGFGFASKTKKKHFVQQKVKVFRAADPLV GVFLWGVAHSINELSQVPPPVMLLPDDFKAS SKIKVNNHLFHRENLPSHFKFKEYCPQVFRNL RDRFGIDDQDYLVSLTRNPPSESEGSDGRFLIS YDRTLVIKEVSSEDIADMHSNLSNYHQVRPLS SPILSLSSLLTYSSAIVSNRCQLGRKLIGRENP
1268	2618	A	10005	2	209	GEGYELFVPSNGVPAVCHMVGRRPHRAVLSP SQDELEHSLGESAAQGAAGVVLWVSWENTR

SEQ ID	SEQ ID	Met	Lero	Predicted	Dunding day	TA
			SEQ		Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1	}	1	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	:	l	i	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	sequence	/=possible nucleotide deletion, \=possible
İ			ļ			
	 	<u> </u>		sequence		nucleotide insertion
						TKVSLGLA
1269	2619	A	10010	245	688	FGMLKNKGHSSKKDNLAVNAVALQDHILHD
1	ļ	i				LQLRNLSVADHSKTQVQKKENKSLKRDTKAI
			}		ŀ	IDTGLKKTTQCPKLEDSEKEYVLDPKPPPLTL
					ļ	AQKLGLIGPPPPPLSSDEWEKVKQRSLLQGDS
			i i		}	
1050	2600	ļ <u>.</u>	1000			VQPCPICKEEFELRPQVFSIRG
1270	2620	A	10011	2.	588	RVDDFVRPLPPGLMSRSRASIHRGSIPAMSYA
[[[PFRDVRGPSTHRTQYVHSPYDRPGWNPRFCII
						SGNQLLMLDEDEIHPLLIRDRRSESSRNKLLR
						RTVSVPVEGRPHGEHEYHLGRSRRKSVPGGK
1						QYSMEGAPAAPFRPSQGFLSRRLKSSIKRTKS
1						QPKLDRTSSFRQILPRFRSADHDRYRGWSMW
1			1			DEIDV
1271	2621		10012	200	262	
12/1	2021	Α	10013	209	363	LPAPPNLSPRLSFGFQFPGGNDNYLTITGPSHP
						FLSGAEVSQSCRRRGGRA
1272	2622	Ā	10014	7	388	SAVTISWKWRSVMGIQTSPALLASLGAGLVT
1	}					LLGLAVGSYLVRRSRRPQVTLLDPNEKDLLR
						LIDKTLSARSPCKHIYLSTRIDGSLSIRPYTPVT
1	1					SDEDQGYVDIDIKVYLKGVHPTFPEGGKMSH
1273	2623	A	10016	ī	1339	
12/3	2023	A	10010	1	1339	MAARTLGRGVGRLLGSLRGLSGQPARPPCGV
1						SAPRRAASGPSGSAPAVAAAAAQPGSYPALS
1		·	1			AQAAREPAAFWGPLARDTLVWDTPYHTVW
						DCDFSTGKIGWFLGGQLNVSVNCLDQHVRKS
1						PESVALIWERDEPGTEVRITYRELLETTCRLA
)]					NTLKRHGVHRGDRVAIYMPVSPLAVAAMLA
						CARIGAVHTVIFAGFSAESLAGRINDAKCKVV
	l					ITFNQGLRGGRVVELKKIVDEAVKHCPTVQH
	1			1		VLVAHRTDNKVHMGDLDVPLEQEMAKEDP
	İ		ľ	ĺ		VCAPESMGSEDMLFMLYTSGSTGMPKGIVHT
1			- 1			QAGYLLYAALTHKLVFDHQPGDIFGCVADIG
			1			WITGHSYVVYGPLCNGATSVLFESTPVYPNA
			l			GRYWETVERLKINQFYGAPTAVRLLLKYGD
1 1			}		,	AWVKKYDRSSLRTLGSVGEPINCEAWEWLH
1						RVVGDSRCTLVDTWWQT
1274	2624	A	10017	1	3750	FRPOGTPRSPASHVLTMSAPDEGRRDPPKPKG
12/4	2024	^	10017	1	3/30	
1 1	' i		i	1	ĺ	KTLGSFFGSLPGFSSARNLVANAHSSARARPA
1						ADPTGAPAAEAAQPQAQVAAHPEQTAPWTE
			į			KELQPSEKMVSGAKDLVCSKMSRAKDAVSS
1 1	1		1			GVASVVDVAKGVVQGGLDTTRSALTGTKEV
						VSSGVTGAMDMAKĠAVQGGLDTSKAVLTG
1						TKDTVSTGLTGAVNVAKGTVQAGVDTTKTV
	i	-		1		LTGTKDTVTTGVMGAVNLAKGTVQTGVETS
i i	Ī		ĺ	Ï	[
1 1	j		i			KAVLTGTKDAVSTGLTGAVNVARGSIQTGV
1 1	1	ſ		ĺ	,	DTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT
1 [I			l		GVDTSKTVLTGTKDTVCSGVTGAMNVAKGT
1 1	ľ	l				IQTGVDTSKTVLTGTKDTVCSGVTGAMNVA
1		1	1	1		KGTIQTGVDTTKTVLTGTKNTVCSGVTGAVN
1		i	ļ	i		LAKEAIQGGLDTTKSMVMGTKDTMSTGLTG
		İ		I		AANVAKGAMOTGLNTTONIATGTKDTVCSG
1		}	1	1		, , , , , , , , , , , , , , , , , , , ,
1		1	- 1	İ		VTGAMNLARGTIQTGVDTTKIVLTGTKDTVC
		l		l		SGVTGAANVAKGAVQGGLDTTKSVLTGTKD
1		1	1	. 1		AVSTGLTGAVNVAKGTVQTGVDTTKTVLTG
1 1		1	1	1	1	TKDTVCSGVTSAVNVAKGAVQGGLDTTKSV
				l		VIGTKDTMSTGLTGAANVAKGAVQTGVDTA
		ļ	- 1	ŀ	1	KTVLTGTKDTVTTGLVGAVNVAKGTVQTGM
			1	J	1	DTTKTVLTGTKDTIYSGVTSAVNVAKGAVOT
		l	-		Į	GLKTTQNIATGTKNTFGSGVTSAVNVAKGAA
!		İ		ł		QTGVDTAKTVLTGTKDTVTTGLMGAVNVAK
1	1	İ	1	1	1	
<u></u>	l					GTVQTSVDTTKTVLTGTKDTVCSGVTGAAN

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isolcucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Scrine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /-possible nucleotide deletion, \=possible nucleotide insertion
						VAKGAIQGGLDTTKSVLTGTKDAVSTGLTGA VKLAKGTVQTGMDTTKTVLTGTKDAVCSGV TGAANVAKGAVQMGVDTAKTVLTGTKDTV CSGVTGAANVAKGAVQTGLKTTQNIATGTK NTLGSGVTGAAKVAKGAVQGGLDTTKSVLT GTKDAVSTGLTGAVNLAKGTVQTGVDT AKTVLSGAKDAVTTGVTGAVNVAKGTVQTGVDT AKTVLSGAKDAVTTGVTGAVNVAKGTVQTG VDASKAVLMGTKDTVFSGVTGAMSMAKGA VQGGLDTTKTVLTGTKDAVSAGLMGSGNVA TGATHTGLSTFQNWLPSTPATSWGGLTSSRT TDNGGEQTALSPQEAPFSGISTPPDVLSVGPEP AWEAAATTKGLATDVATFTQGAAPGREDTG LLATTHGPEEAPRLAMLQNELEGLGDIFHPM NAEEQAQLAASQPGPKVLSAEQGSYFVRLGD LGPSFRQRAFEHAVSHLQHGQFQARDTLAQL QDCFRL
1275	2625	A	10025	124	415	TILARKKEKTCPCKKEIGRNSRSGMYSRKAM YKRKYSAANTKVEKKKKEKVLAPVTKPVGG DKNGGTRVVKLPTMPRYYPTEDVPRKLLSHG KKPFS
1276	2626	A	10030	3	507	GGSLRFSPPRVPSCSRVFCPVPPGGCGLPSPMS ASRPQSPTTPWCLPRRYMKHKRDDGPEKQED EAVDVTPVMTCVFVVMCCSMLVLLYYFYDL LVYVVIGIFCLASATGLYSCLAPCVRRLPFGK CRIPNNSLPYFHKRPQARMLLLALFCVAVSV VWGVFRNEDO
1277	2627	A	10035	51	869	YSRTYPLPATMASSEVARHLLFQSHMATKT TCMSSQGSDDEQIKRENIRSLTMSGHVGFESL PDQLVNRSIQQGFCFNILCVGETGIGKSTLIDT LFNTNFEDYESSHFCPNVKLKAQTYELQESN
						VQLKLTIVNTVGFGDQINKEERQLGRSQSTEN PQKYRSEQHPVEPKKCTSFWKGALGKWAGIE SSGQSAQQPYLPINSPPHRLADVADVHLFSSV LSGAFGCYHLDVTVNEFKKQQNRDEQEGYS KGDQEQGSWKHGADPLRGGEM
1278	2628	A	10036	3	457	RAFDVRRKKSLRPCCPRDFHAGCLTVSGPST VMGAVGESLSVQCRYEEKYKTFNKYWCRQP CLPIWHEMVETGGSEGVVRSDQVIITDHPGDL TFTVTLENLTADDAGKYRCGIATILQEDGLSG FLPDPFFQVQVLVSSASSTENSVKTP
1279	2629	A	10039	214	435	NDSLVPMSSWRSCARAPSSESAWRRSAATRR SRKCLRTKRKRWSSGKGTQMQSTLSETPRRA QMPCMWWYPFWG
1280	2630	A	10043	2	344	RATWHNAGKEREAVQLMAGAEKRVKASHS FLRGLFGGNTRIEEACEMYTRAANMFKMAK NWSAAGNAFCQAAKLHMQLQSKHDSATSFV DAGNAYKKADPQGKTARHVACYLCV
1281	2631	A	10080	620	818	VIYKLDSSLFSYFIYFFIFETESHFLPLMKWTG PIMAHCSLKILASRNSADSAFLSAGDTSLSHST
1282	2632	A	10084	3	1640	SASIIIRGDKRASGEVGIAPSSRHILIGEPSAKY NGTAIISLVRGPGILGEVTVFWRIFPPSVGEFA ETSGKLTMRDEQSAVIVVIQALNDDIPEEKSF YEFQLTAVSEGGVLSESSSTANITVVASDSPY GRFAFSHEQLRVSEAQRVNITIIRSSGDFGHVR LWYKTMSGTAEAGLDFVPAAGELLFEAGEM RKSLHVEILDDDYPEGPEEFSLTITKVELQGR GYDFTIQENGLQIDQPPEIGNISIVRIIIMKNDN AEGIIEFDPKYTAFEVEEDVGLIMIPVVRLHGT

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutaminc, R=Argininc, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion YGYVTADFISQSSSASPGGVDYILHGSTVTFQ HGQNLSFINISIIDDNESEFEEPIEILLTGATGG AVLGRHLVSRIIIAKSDSPFGVIRFLNQSKISIA NPNSTMILSLVLERTGGLLGEIQVNWETVGPN SQEALLPQNRDIADPVSGLFYFGEGEGGVRTII LTIYPHEEIEVEETFIIKLHLVKGEAKLDSRAK DVTLTIQEFGDPNGVVQFAPETLSKKTYSEPL ALEGPLLITFFVRRVKGTFGEIM
1283	2633	A	10088	316	516	MGSKTLPAPVPIHPSLQLTNYSFLQAVNGLPT VPSDHLPNLYGFSALHAVHLHQWTLGYPAM HLXRS
1284	2634	A	10091	2	569	FVSPSRAMASALIYVSKFKSFVILVVTPLLLLP LVILMPAKFVRCAYVIILMAIYWCTEVIPLAV TSLMPVLLFPLFQILDSRQVCVQYMKDTNML FLGGLIVAVAVERWNLHKRIALRTLLWVGA KPARLMLGFMGVTALLSMWISNTATTAMMV PIVEAILQQMEATSAATEAGLELVDKGKAKE LP
1285	2635	A	10092	290	728	KQSTRPDVMTLYPLHWQEEMSGESVVSSAVP AAATRTTSFKGTSPSSKYVKLNVGGALYYTT MQTLTKQDTMLKAMFSGRMEVLTDSEGWIL IDRCGKHFGTILNYLRDGAVPLPESRREIEELL AEAKYYLVQGLVEECQAALQV
1286	2636	Α	10100	1	574	RPRGRGAWAGPGGDYSGVRRQQRRRTRISGS QRGSDAAGTMGCCTGRCSLICLCALQLVSAL ERQIFDFLGFQWAPILGNFLHIIVVILGLFGTIQ YRPRYIMVYTVWTALWVTWNVFIICFYLEVG GLSKDTDLMTFNISVHRSWWREHGPGCVRR VLPPSAHGMMDDYTYVSVTGCIVDFQYLEVI HSA
1287	2637	A	10103	252	376	RSRMGDKPIWEQIGSSFIQHYYQLFDNDRTQL GAIYVSFQL
1288	2638	A	10107	1	478	MEEEDESRGKTEESGEDRGDGPPDRDPTLSPS AFILRAIQQAVGSSLQGDLPNDKDGSRCHGL RWRRCRSPRSEPRSQESGGTDTATVLDMATD SFLAGLVSVLDPPDTWVPSRLDLRPGESEDM LELVAEVRIGDRDPIPLPVPSLLPRLRAWRTG KT
1289	2639	A	10113	237	438	LLSRMPSTNRAGSLKDPEIAELFFKEDPEKLFT DLREIGHGSFGAAYFARDVRTNEVVAIKKMS YSG
1290	2640	Α	10114	367	856	RGAKAKSAVLPPGPPCSSILILSPPAPLTPRSPG TEATRPTAMSKSLKKKSHWTSKVHESVIGRN PEGQLGFELKGGAENGQFPYLGEVKPGKVAY ESGSKLVSEELLLEVNETPVAGLTIRDVLAVI KHCKDPLRLKCVKQGESSGLLSVLPGGGTAR GAGQ
1291	2641	A	10116	128	591	RTIRETERRSALSCSVLKSEPLPGLQPQASQQR RRRLPGRRQVQVQEGGGSGLRAWVLAMASV LGSGRGSGGLSSQLKCKSKRRRRRSKRKDK VSILSTFLAPFKHLSPGITNTEDDDTLSTSSAE VKENRNVGNLAARPPPSGDRARGGATR
1292	2642	A	10121	1	749	QRRRFRAGLWGGHGLTDGLRRNGGCGCSAR VPRVGERLRGHRCPDPLCLLLDMLFLSFHAG SWESWCCCCLIPADRPWDRGQHWQLEMADT RSVHETRFEAAVKVIQSLPKNGSFQPTNEMM LKFYSFYKQATEGPCKLSRPGFWDPIGRYKW DAWSSLGDMTKEEAMIAYVEEMKKIIETMP MTEKVEELLRVIGPFYEIVEDKKSGRSSDITSD

No. of N	SEQ ID	SEO ID	Met	SEC	Predicted	Tr. 4: 4 4	
nuche cotide scipe contest of the co				SEQ		Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Description Description	_		, nou				D=Aspartic Acid, E=Glutamic Acid,
Section		,]		1		r=rnenylalanine, G=Glycine, H=Histidine,
1293 2643 A 10124 2 989 PLMSLVRVVEFVAASSAQKTFSRLENYTMVC KADEKFNOLVHEIR, WFFYSSGL CRGRIGRSALVFLENKER, CRGRIFF RALED FOR STANDARD CREATER STANDARD	1						M-Methioning N-Assessing D-D-line
mino acid recidius of peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence per control of the sequence per con				1			O=Glutamine R=Arginine S=Serine
Personal Personal		ł]	***			T=Threonine V=Valine W=Tryptophan
Populac		ŀ					Y=Tyrosine X=I Inknown *=Ston coden
1293 2643 A 10124 2 989 PILMSLYRVPKAASSAQKTPSRLENYYMVC			1	ł		1 1 1	/=possible nucleotide deletion \=possible
LGNVLTSTPAKTVNGKAESSDSGAESEEEE	1					1	nucleotide insertion
1293 2643 A 10124 2 989 PI_MSI_VR_VVEFVAASSAQKTPSRLENYYMVC				1	ļ -		
						1	
CGRGIRDSARMCSTCACVEYYGKALEVLYGGULV	1293	2643	A	10124	2	989	PLMSLVRVVEFVAASSAQKTPSRLENYYMVC
CGRGIRDSARMCSTCACVEYYGKALEVLYGGULV			i				KADEKFNQLVHFLRNHKQEKHLVFFRYSSGL
CTDVMARGIDIPEVNWU_QVDPPSNASAFVH RCGRTARIGHGSAL VFLIPMEXSMALADS VERGMARFVSYVNQ-AVARFECH_FRILADI DFASLARGFALLEN/PKLKSMALADS VFREGMKAFVSYVQ-AVARFECH_FRILADI DFASLARGFALLEN/PKLMPELROK PFPLADI DFASLARGFALLEN/PKLMPELROK PFPLADI DFASLARGFALLEN/PKLMPELROK PFPLADI DFASLARGFALLEN/PKLMPELROK PFPLADI DFASLARGFALLEN/PKLMPELROK PFPLADI DFASLARGFALLEN/PKLMPELROK PFPLADI DFASLARGFALLEN/PKLMPELROK PFPLADI DFASLARGFALLEN/PKLMPELROK PFPLADI DFASLARGFALLEN/PKLMPELROK PFPLADI DFASLARGFALLEN/PKLMPELROK PFPLADI DFASLARGFALLEN/PKLMPELROK PFPLADI DFASLARGFALLEN/PKLMPELROK PFPLADI DFASLARGFALLEN/PKLMPELROK PFPLADI PFPLAD			1			1	CGRGIRDSARMCSTCACVEYYGKALEVLVK
RCGRTARIGHGGSAL VFILPMEESYINLAIN QRCPLQEMKCPRNTNADLIPKILSMALADRA VFEKGMKAPVSYVQAYAKHECNIJERILKDIL DFASLARGEALILMPKMPHEKOKOPPDFVVV DVNTDTIPFKDKIBEKORQKILLEQQRREKTEN EGRRKFIKNKAWSKQKAKKK VTMYRDCIESTGDYFILLDABGPWGIILESLA ILGIVVTILLILAFLFIJMRKIQDCSQWNUPTQ LLFILSVLGILGAFAPIRIELNQQTAPVRYFIF_GVLFALCFSGLLAHASNLVKLVRGCVSSWT TILCIAIGCSLQIIIATEYVTLMTRGMMFVN MTPCQLNVDFVVLLVYVJRIMALITFVSKAT FCGPCENWKOHGRIJFITVLYMSKAM LRGNPGPQROPOWDDPVCLALVTNAWVFL LLYVPELCILVJRSCRQECPJOHACPVTAYQ HSFQVENQELSRDKWKVLINSDFISHSGA GRGAPRQEGGSSWRQV HSFQVENQELSRDKWKVLINSDFISHSGA GRGAPRQEGGSSWRQV SWSLDPFMGIMSGQVGDLSPSQEKSLAQFE NIQDVLSALPNPDDYFLLRWQARSPDLQKS EDMLRHMEFRKQQDLANILAWQPFVVRL YNANCIGHIDEGSSPWWITVGSQDPKGLLL SASKQELLRDSFRSCELLRECELQSQKLGKR VEKINFREGGSPWSTIVGSQDPKGLLL SASKQELLRDSFRSCELLRECELQSQKLGKR VEKINFREGGSPWSTIVGSQDPKGLLL SASKQELLRDSFRSCELLRECELQSQKLGKR VEKINFREGGSPWSTIVGSQDPKGLL SASKQELRDSFRSCELLRECELQSQKLGKR VEKINFREGGSPWSTIVGSQDPKGLL SASKQELRDSFRSCELLRECELQSQKLGKR VEKINFREGGSPWSTIVGSQDPKGLL SASKQELRDSFRSCELLRECELQSQKLGKR VEKINFREGGSPWSTIVGSQDPKGLL SASKQELRDSFRSCELLRECELQSQKLGKR VEKINFREGGSPWSTIVGSQDPKGLL SASKQELRDSFRSCELLRECELQSQKLGKR VEKINFREGGSPWSTIVGSQDPKGLL SASKQELRDSFRSCELLRECELQSQKLGKR VEKINFREGGSPWSTIVGSQDPKGLL SASKQELRDSFRSCECTPQCC QFVCCQFTCCRFSCCETTCCHPXCC QCMYLRPDQTTRRSFYRMSRLARHQLVTKI QQEMYLRPDQTTRRSFYRMSRLARHQLVTKI QQEMYLRPDQTTRRSFYRMSRLARHQLVTKI QQEMYLRPDQTTRRSFYRMSRLARHQLVTKI QQEMYLRPDQTTRRSFYRMSRLARHQLVTKI QQEMYLRPDQTTRRSFYRMSRLARHQLVTKI QQEMYLRPDQTTRRSFYRMSRLARHQLVTKI QQEMYLRPDQTTRSFYRMSRLARHQLVTKI QQEMYLRPDQTTRSFYRMSRLARHQLVTKI QCEMALRAGRORGVSARTGGSKATTRSGSFT SASQLEMBEAPKR VSLALQI PEHGSKDIGN VPGNCSSIPPQNGGTAQTA SAGLEMBEAPKR VSLALQI PEHGSKDIGN VPGNCSSIPPQNGGTAQTA SAGLEMBEAPKR VSLALQI PEHGSKDIGN VPGNCSSIPPQNGGTAQT SAGLEMBEAPKR VSLALQI PEHGSKDIGN VPGNCSSIPPQNGGTAQTA VTTGILGPGLIGNILALWYYGYMKERKA VIFMNLARHQLVALURARH VTTGRGGRASSSISDLQGKGFGKGTA VTTGRI VTTGRGGGRASSSISDLQGKGFGKGTA VTTGRI							
	ĺ		ĺ	i		l	CTDVMARGIDIPEVNWVLQYDPPSNASAFVH
1294							RCGRTARIGHGGSALVFLLPMEESYINFLAIN
DPASLARGFALLRMPKMPELRGKGPPDFVPV	1			ì			QKCPLQEMKPQRNTADLLPKLKSMALADRA
DVNTDTIPFRDKIREKOR(KLLEQORTEKTEN	i		ľ	l	}	1	VFEKGMKAFVSYVQAYAKHECNLIFRLKDL
							DFASLARGFALLRMPKMPELRGKQFPDFVPV
1294							
ILGIVVTILLIAPLEPIMRIQDCSQWNVLPTQ LIFILSVLGLFGLAFAFIIELNQQTMYRYFLF GVI-FALCFSCILAHASHUKLVRGCVSFSWT TILCIAIGCSLLQIIIATEYVTLIMTRGMMFVAN MPTPCQLNVDFVVLLVVDFVVLLVVTLMALTFFVSKAT FCGPCENWKQHGRLIFITVLFSIIIWVVWISML LRQNPQFQRQPWDDPVVCIALVTNAWVFIL LLYIVPELCILYRSCRQELQGNACPVTAYQ HSFQVENQELSRDKWKVLINSDFLSISGA GRAPRQEOPGSSWRQV HSFQVENQELSRDKWKVLINSDFLSISGA GRAPRQEOPGSSWRQV GRAPRQEOPGSSWRQV SEWSLDPFMGIMSGQVGDLSPSQEKSLAQFRE NIQDVLSALIPPDDYFLRWLQARSFDLQKS EDMLRKHMEFRRQQDLANILAWQPPEVVRL YANGIGCHDGEGSPWHYGSQDPFGGLL SASKQELLRDSFRSCELLIRECELGSQKLGKR VERHAMFIGEGGLRDKWFGIELLQS GRAPRQEOPGSSWRQV GRAPRQEOPGSSWRQV SASKQELLRDSFRSCELLIRECELGSQKLGKR VERHAMFIGEGGLRDKWFGIELLQS GRAPRQEOPGSCETTCCHPSCCQSVCCQPT TTCRTTCCRPSCCVSSCCRPQCCQDVCCQPT SASKQELLRDSFRSCELLIRECELGSQKLGKR VERHAMFIGEGGLRDKWFGIELLQS GRAPRQEABASHSQVPRTPASGCYVLNSMTPEG GEMYLRFDOTTRRSPYRMSRLARHQLVTKI QQEIEAKEACDWLRAAGFPQYAQLYEDSQFP GRAPRQEABASHSQVPRTPASGCYVLDSDGFP GRAPRQEABASHSQVPRTPASGCYVLDSDGFP GRAPRQEABASHSQVPRTPASGCYVLDSDGFP GRAPCALACKVSRPCTRLFSTEXAFPVVEGG GRAPRQEABASHSQVPRTPASGCYVLDSDGFP KGRCELACKVSRPCTRLFSTEXAFPVVEGG VCHHV VFYGMKETKRA VIFMINLAHDLLQVLSLPLRIFYYLKHDWPF VFY	1204	2644		10100		10.40	
LLFLLSVLGLEGLAFAFIELNQOTAPVRYPLF GVLFALCFSCLLAHACTSCLAHANLYLVRGCYSFWT TILCIAIGCSLLQIIIATEYVTLIMTRGMMFVN MTPCQLNVDFVVLLVYVLFMALTFFVSKAT FCGPCENWKQHGRLIPITVLFSIIIWVVWISML LRGNPQFQRQPQWDDFVVCLAVTNAWVFL LLYIVPELCILYRSCRQECPLQGNACPVTAYQ HSFQVENQELSRDKWKVLLNSDFLSHSGA RPRVVTHNSQWCFLPQDHPGWLFQGSGAPG GRGAPRQEGFGSSWRQV GRGAPRQEGFGSSWRQV LSWSLDPFMGIMSGQVGDLSFSDEKSLAGPRE MIQDVLSALPPRDDYFLRWLQARSFDLQKS EWSLDPFMGIMSGQVGDLSFSDEKSLAGPRE MIQDVLSALPPRDDYFLRWLQARSFDLQKS EWSLDPFMGIMSGQVGDLSFSDEKSLAGPRE MIQDVLSALPRIPDLYFLRWLQARSFDLQKS EWSLDPFMGIMSGQVGDLSFSDEKSLAGPRE MIQDVLSALPRIPDLYFLRWLQARSFDLQKS EWSLDPFMGIMSGQVGDLSFSDEKSLAGPRE MIQDVLSALPRIPDLYFLRWLQARSFDLQKS EWSLDPFMGIMSGQVGDLSFSDEKSLAGPRE A 10138 48 407 MVSSCGSVCSDQGCGQDLCQETCCRPSCCE TTCRRTCCRPSCCVSSCCRPQCCQSVCCQPT CSRPSCCQTTCCRTTCXPSCCSCCRQCCQVCCQPT CSRPSCCQTTCCRTTCXPSCCSCCRQCCQVCCQPT CSRPSCCQTTCCRTTCXPSCCSCCRQCCQVCCQPT CSRPSCCQTTCCRTTCXPSCCSCCRQCCQVCCQPT CSRPSCCQTTCCRTTCXPSCCSCCRQCCQVCCQPT CSRPSCCQTTCCRTTCXPSCCSCCRQCCQVCCQPT CSRPSCCQTTCCRTTCXPSCCSCCRQCCQVCCQPT CSRPSCCQTTCCRTTCXPSCCSTCTCCPDCC CSRPSCCQTTCCRTTCXPSCCSTCTCCPDCC CSRPSCCQTTCCRTTCXPSCCSTCTCPDCC CSRPSCCQTTCCRTTCXPSCCSTCTCPDCCT CSRPSCCQTTCCRTTCXPSCCSTCTCPDCC CSRPSCCQTTCCRTTCXPSCCSTCTCPDCC CSRPSCCQTTCCRTTCXPSCCSTCTCPDCC CSRPSCCQTTCCRTTCXPSCCSTCTCPDCCT CSRPSCCQTTCCRTTCXPSCCTTCCPDCCT CSRPSCCQTTCCRTTCXPSCCTTCCTTCTTCTTTCTTCTTTCTTTCTTTCTTTTCT	1294	2044	A	10129	91	1042	
GVLFALCFSCLLAHASNILVKLVRGCVSFSWT TILCIAIGCSLIQUIALTEYVTLIMTRGMMFVN MTPCQLNVDFVVLLVYVLFMALTFEVSKAT FCGPCENWKQHGRLIFITVLFSIIWVVWISML LRQNPpCFQRQPQWDVPVCIALVTNAWVFL LLYIVPELCILYRSCRQECPLQGNACPVTAYQ HSFQVENQELSRDKWKVLINSDFLSHSGA RPRVVTHNSQWCELPQDHPGWLFGQSGAPG GRQAPRQEGFGSSWRQV GRQAPRQEGFGSSWRQV EWSLIPPMGMMSGQVGDLSPSQEKSLAQFRE NIQDVLSALPNPDDYFLLRWLQARSFDLQKS EDMLRKIMEFRKQDLANILAWQPPEVVRL YANAIGIGHDGEGSVWHVHQSQDPFVGLL SASKQELLRDSFRSCELLLRECELQSQKLGKR VEKHAFGLEGLGLRDLWKPGHELLQE T297		Ĭ	1	1	1	İ	ILGIVVTILLLLAFLFLMRKIQDCSQWNVLPTQ
TILCIAIGCSLIQIIIATEYVTLIMTROMMFVN MTPCQLNVDFVVLLVYVLEMALTFFVSKAT FCGPCENWKQHGRLIPITVLFSIIWVVWISML LRGNPQFQRQPQWDDPVVCLIAVTNAWVFIL LLYIVPELCILYRSCRQECPLQGNACPVTAYQ HSFQVENQELSRDKWKVLINSDFLSHSGA RPRVVTHNSQWCFLPQDHPGWLFGQSGAPG GRGAPRQEGFGSSWRQV GRGAPRQEGFGSSWRQV HSFQVENQELSRDKWKVLINSDFLSHSGA RPRVVTHNSQWCFLPQDHPGWLFGQSGAPG GRGAPRQEGFGSSWRQV GRGAPRQEGFGSSWRQV HSFQVENQELSALPPRDVFLLRWQARSPILQKS EDWLRKHMEFRKQQLANILAWQPPEVVRL YNANGICGHDEGSPVWYHLVQSQDPRGLIQ SASKQELLRDSFRSCELLLRECELQSQKLGKR VPEKIMFGLEGEGSPWYHYHVGSQDPRGLIQ SASKQELLRDSFRSCELLLRECELQSQKLGKR VPEKIMFGLEGEGGFGGGDLOGFTCGRFSCCE TTCCRTTCCRPSCCVSSCCRPQCC TTCCRTTCCRPSCCVSSCCRPQCC TTCCRTTCCRTCCRTTCCRPSCCVSSCCRPQCC TTCCRTTCCRTTCCRTSCVSSCCRPQCC GPVCCQPTCCRPSCCETTCCHPXCC GGMRKSAEMFSQVPRTPASGCYTLNSMTPEG GGMRKSAEMFSQVPRTPASGCYTLNSMTPEG GEMYLRPQTTRRSPYRMSILARHQLVTKI QZEIEAKEACDWLRAAGFPQYAQLYEDSGPP NIVAVKNDHDFLEKDLGEPLCRRINT ASAQLENMEEAPKRVSLALQLPEHGSKDIGN VPGNCSENPCQNGGTCVPGADAHSCDCGPG KGRCELACIKVSRCTRLFSETKAFPVWEGG VCHHV ASAQLENMEEAPKRVSLALQLPEHGSKDIGN VPGNCSENPCQNGGTCVPGADAHSCDCGPG KGRCELACIKVSRCTRLFSETKAFPVWEGG VCHHV AKIASLERIMPANYTCTRPDGDNTDFRYFTYA VTYTGILQPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV TTTGILQPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV TTTGILQPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV TTTGILQPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF PGRRVGITSOTGLSSNLOENCSKLAFISSHTE KCQLQCMPMEGRGRASSISDLOGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWIJGKPDFGRRRRGSPGEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSISSLHVYDENKNISKIERHKRRTSTPV ARRISESLHVYDENKNISKIERHKRRTSTPV ARRISESLHVYDENKNISKIERHKRRTSTPV ARRISESLHVYDENKNISKIERHKRRTSTPV ARRISESLHVYDENKNISKIERHKRRTSTPV ARRISESLHVYDENKNISKIERHKRRTSTPV ARRISESLHVYDENKNISKIERHKRRTSTPV ARRISESLHVYDENKNISKIERHKRRTSTPV ARRISESLHVYDENKNISKIERHKRRTSTPV ARRISESLHVYDENKNISKIERHKRRTSTPV ARRISESLHVYDENK					:		
MTPCQLNVDFVVLLVYLLMALTIFVSKAT FCGPCEW KQhRGLIFITVLFSIIWVVWISML LRGNPQFQRQPQWDDPVVCIALVTNAWVFL LLTIVPELCILYRSCRQECPLQGNACPVTAYQ HSFQVENQELSDEW KVLLNSDFLSHSGA HSPQVVENQELSDEW KVLLNSDFLSHSGA HSPQVVENQELSDEW KVLLNSDFLSHSGA HSPQVVENDEW KYLLNSDFLSHSGA HSPQVVENDEW KYLLNSDFLSHSGA HSPQVVENDEW KYLLNSDFLSHSGA GRGAPRQEGPGSSWROP KYLLNSDFLSHSGA GRGAPRQEGPGSSWROP KOLL SASKQELLRDSFRSCELLRECLQSQKLGKR HQDVLSALPNPDDYFLLRECLQSQKLGKR HQDVLSALPNPDDYFLLRECLQSQKLGKR HQDVLSALPNPDDYFLLRECLQSQKLGKR HQDVLSALPNPDGYFLLRECLQSQKLGKR HQDVLSALPNPDGYFLLRECLQSQKLGKR HQDVLSALPNPDGYFLRECLQSQKLGKR HQDVLSALPNPDGYFLRECLQSQKLGKR HQDVLSALPNPDGYFLARCELQSQKLGKR HQDVLSALPNPDGYFLARCELQSQKLGKR HQDVLSALPNPDGYFLARCELQSQKLGKR HQDVLSALPNPDGYFLARCELQSQKLGKR HQDVLSALPNPDGYFLARCELQSQKLGKR HQDVLSALPNPDGYFLARCELQSQKLGKR HQDVLSALPNPDGYFLARCELQSQKLGKR HQDVLSALPNPDGYFLARCELQSQKLGKR HQDVLSALPNPDGYFLARCELQSQKLGKR HQDVLSALPNPDGYFLARCELQSQKLGKR HQDVLSALPNPDGYFLARCELQSQKLGKR HQDVLSALPNPDGYFLARCELQSQKLGKR HQDVLSALPNPDGYFLARCELQSQKLGKR HQDVLSALPNPDGYFLARCELQSQKLGKR HQDVLSALPNPDGYFLARCELQSQKLGKR HQDVLSALPNPDGYFLARCELQSQKLGKR HQDVLSALPNPDGYFLARCELQSQKLGKR HDDVLSALPNPDGYFLARCELQSQKLGKR HQDVLSALPNPDGYFLARCELQSQKLGKR HDDVLSALPNPDGYFLARCELQSQKLGKR HDDVLSALPNPDGYFLARCELQSQKLGKR HDDVLSALPNPDGYFLARCELQSQKLGKR HDDVLSALPNPDGYFLARCELQSQKLGKR HDDVLSALPNPDGYFLARCELQSQKLGKR HDDVLSALPNPDGYFLARCELQSGK HDDVLSALPNPDGYFLARCELQSGK HDDVLSALPNPDGYFLARCELQSGK HDDVLSALPNPDGYFLARCELQSGK HDDVLSALPNPDGYFLARCELQSGK HDDVLSALPNPDGYFLARCELQSGK HDDVLSALPNPDGYFLARCELQSGK HDDVLSALPNPDGYFLARCELQSGK HDDVLSALPNPDGYFLARCELQSGK HDDVLSALPNPDGYFLARCELQSGK HDDVLSALPNPDGYFLARCELQSGK HDDVLSALPSK	1	Ī	1				
FCGPCENWKQHGRLIFITVLFSIIIWVVWISML	l l	ı			1	}	TILCIAIGCSLLQIIIATEYVTLIMTRGMMFVN
LRGNPQF0RQPQWDDPVVCIALVTNAWVFL	İ				I		MIPCQLNVDFVVLLVYVLFLMALTFFVSKAT
LLYIVPELCIL YRSCROECPLOGNACPYTAYO					1		
HSFQVENQELSRDK WK VLINSDFLSHSGA		1	1		1	}	
1296	Ì						
1296	1295	2645	 	10122	276	510	
1296	12/3	2043	^	10133	370	219	
NIQDVLSALPNPDDYFLLRWLQARSFDLQKS EDMLRKHMEFRKQQDLANILAWQPPEVVRL YNANGICGHDGEGSPVWYHIVGSQDPKGLLL SASKQELLRDSFRSCELLLRECELQSQKI.GKR VEKHAIFGLEGLGLRDLWKPGIELLQE 1297 2647 A 10138 48 407 MVSSCCGSVCSDQGCGQDLCQETCCRFSCCE TTCCRTTCCRTSCCVSSCCRPQCCQSVCCQPT CSRPSCCQTTCCRTTCYPRSCCVSSCCRPQCCQ QPVCCQPTCCRPSCCETTCCHPXCC QPVCQPTCCRPSCCETTCCHPXCC QPVCQPTCCRPSCCETTCCHPXCC QPWCLRFPQTTRRSPYRMSRILARHQLVTKI QQEIEAKEACDWLRAAGFPQYAQLYEDSQFP INIVAVKNDHDFLEKDLGEFLCRLINT PRFSELVDGRGRVSARFGGSPSKAATVRSQPT ASAQLENMEEAPKRVSLALQLPEHGSKDIGN VPGNCSENPCQNGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV 1300 2650 A 10162 98 391 AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV 1301 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFKGTG EKNPGVGSARRISPQASAGGSPWQRGAQT RWI.GKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNOLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNFSKIEREHKRRTSTPV	1296	2646	Α	10135	2	551	
EDMLRKHMEFRK QQDLANILAWQPPEVVRL	1270	1 -0.0	١	10155	3	331	MICONI SAL BARDONELL BARLOY BEEDLOKE
1297 2647 A 10138 48 407 WSSCCGSVCSDQGCQDLCQETCCRPSCCE TTCCRTTCCRPSCCVSCCRPQCCQSVCCOPT CSRPSCCQTTCCRTSCCPCCQSVCCQPT CSRPSCCQTTCCRTSCCPCCQPVCCQPTCCRPSCCETTCCRPSCCSTCTCCPSCCPQCCQPVCCQPTCCRPSCCETTCCRPSCCSCRPQCCQPVCCQPTCCRPSCCETTCCRPSCCVSSCCRPQCCQPVCCQPTCCRPSCCETTCCPSCCPQCCQPVCCQPTCCRPSCCETTCCPSCCPQCCQPVCCQPTCCRPSCCETTCCPSCCPQCCQPVCCQPTCCRPSCCETTCCPSCCPQCCQPVCCQPVCCQPTCCRPSCCETTCCPSCCPQCCQPVCCQPTCCRPSCCETTCCPSCCPQCCQPVCCQPTCCRPSCCETTCCPSCCPQCCQPVCCQPTCCRPSCCETTCCPPSCCPQCCQPVCCQPTCCRPSCCETTCCPPSCCPQCCQPVCCQPTCCRPSCCPPCCCQPVCCQPVCCQPTCCRPSCCPPCCCQPCCQPVCCQPTCCRPSCCPTCCPPSCCPQCCQPCCQPVCCQPTCCRPSCCPTCCPPSCCPPCCQPCCQPVCCQPTCCRPSCCPTCCPPSCCPPCCQPCCQPVCCQPVCCQPCCQPVCCQPCCQPVCCQPCCQP				1			
SASKQELLRDSFRSCELLLRECELQSQKLGKR	Į.	[ĺ			YNANGICGHDGEGSPVWYHIVGSODPKGI I I
1297 2647 A 10138 48 407 MVSSCGSVCSDQGGQDLCQETCRPSCCE TTCCRTTCCRPSCCVSSCCRPQCCQSVCCQPT CSRPSCCQTTCCRTTCYRPSCCVSSCCRPQCCQSVCCQPT CSRPSCCQTTCCRTTCYRPSCCVSSCCRPQCCQPVCCQPTCRPSCCETTCCHPXCC QPVCQCMPMEGRGRASGGSPKCATTCATCATCATCATCATCATCATCATCATCATCATCAT					ļ		SASKOFI LEDSERSCELLI RECELOSORI CVD
1297 2647 A 10138 48 407 MVSSCCGSVCSDQGCGQDLCQETCCRPSCCE TTCCRTTCCRPSCCVSCCRPQCCQSVCCQPT CSRPSCCVTTCCRTTCYRPSCCVSCCRPQCC QPVCCQPTCCRPSCCETTCCHPXCC QPVCCQPTCCRPSCCETTCCHPXCC QPVCCQPTCCRPSCCETTCCHPXCC QPVCCQPTCCRPSCCETTCCHPXCC QPVCCQPTCCRPSCCETTCCHPXCC QPVCCQPTCCRPSCCETTCCHPXCC QPVCCQPTCCRPSCCETTCCHPXCC QPVCCQPTCRPSCCETTCCHPXCC QPVCCQPTCRPSCCETTCCHPXCC QEMYLRFDQTTRRSPYRMSRILARHQLVTKI QQEIEAKEACDWLRAAGFPQYAQLYEDSQFP NIVAVKNDHDFLEKDLGEPLCRRLNT ASAQLENMEEAPKRVSLALQLPEHGSKDIGN VPGNCSENPCQNGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV 393 AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV QV VPV 1301 AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV PV PV PV PV PV PV PV PV PV PV PV PV				 			
1298 2648 A 10156 94 453 GGNRKSAEMFSQVPRTPASGCYVINSMTPEG QEMYLRFDQTTRRSPYRMSRILARHQLVTKI QQEIEAKEACDWLRAAGFPQYAQLYEDSQFP NIVAVKNDHDFLEKDLGEPLCRRLNT ASAQLENMEEAPKRVSLALQLPEHGSKDIGN VPGNCSENPCQNGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV 1300 2650 A 10162 98 391 AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV 1301 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWIGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKPA ARRLSESLHVVDENKNFSKIEREHKRRTSTFV	1297	2647	A	10138	48	407	
CSRPSCQTTCCRTTCYRPSCCVSSCCRPQCC					:		TTCCRTTCCRPSCCVSSCCRPOCCOSVCCOPT
1298 2648 A 10156 94 453 GGNRKSAEMFSQVPRTPASGCYYLNSMTPEG QEMYLRFDQTTRRSPYRMSRILARHQLVTKI QEMYLRFDQTTRRSPYRMSRILARHQLVTKI QEIEAKEACDWLRAAGFPQYAQLYEDSQFP INIVAVKNDHDFLEKDLGEPLCRRLNT 1299 2649 A 10161 I 393 PRFSELVDGRGRVSARFGGSPSKAATVRSQPT ASAQLENMEEAPKRVSLALQLPEHGSKDIGN VPGNCSENPCQNGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV 1300 2650 A 10162 98 391 AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV 1301 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWI.GKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNFSKIEREHKRRTSTPV							CSRPSCCOTTCCRTTCYRPSCCVSSCCRPOCC
1298 2648 A 10156 94 453 GGNRKSAEMFSQVPRTPASGCYYLNSMTPEG QEMYLRFDQTTRRSPYRMSRILARHQLVTKI QEEIEAKEACDWLRAAGFPQYAQLYEDSQFP INIVAVKNDHDFLEKDLGEPLCRLNT 1299 2649 A 10161 I 393 PRFSELVDGRGRVSARFGGSPSKAATVRSQPT ASAQLENMEEAPKRVSLALQLPEHGSKDIGN VPGNCSENPCQNGGTCVPGADAHSCDCGPGF KGRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV 1301 2651 A 10165 I 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWI.GKPDPGRKRRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNFSKIEREHKRRTSTPV		<u></u>	i.	1			OPVCCOPTCCRPSCCETTCCHPXCC
CEMYLRFDQTTRRSPYRMSRILARHQLVTKI QQEIEAKEACDWLRAAGFPQYAQLYEDSQFP INIVAVKNDHDFLEKDLGEFLCRRINT 1299 2649 A 10161 1 393 PRFSELVDGRGRVSARFGGSPSKATVRSQPT ASAQLENMEEAPKRVSLALQLPEHGSKDIGN VPGNCSENPCQNGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV 1300 2650 A 10162 98 391 AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV 1301 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV	1298	2648	A	10156	94	453	GGNRKSAEMFSOVPRTPASGCYYLNSMTPEG
I299 2649 A 10161 I 393 PRFSELVDGRGRVSARFGGSPSKAATVRSQPT ASAQLENMEAPKRVSLALQLPEHGSKDIGN VPGNCSENPCQNGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV I300 2650 A 10162 98 391 AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV I301 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV							QEMYLRFDOTTRRSPYRMSRILARHOLVTKI
INIVAVKNDHDFLEKDLGEPLCRRLNT 1299 2649 A 10161 1 393 PRFSELVDGRGRVSARFGGSPSKAATVRSQPT ASAQLENMEEAPKRVSLALQLPEHGSKDIGN VPGNCSENPCQNGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV	1		ĺ				QQEIEAKEACDWLRAAGFPOYAOLYEDSOFP
1299 2649 A 10161 I 393 PRFSELVDGRGRVSARFGGSPSKAATVRSQPT ASAQLENMEEAPKRVSLALQLPEHGSKDIGN VPGNCSENPCQNGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV 1300 2650 A 10162 98 391 AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV 1301 2651 A 10165 I 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWI.GKPDPGRKRRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNFSKIEREHKRRTSTPV			<u></u> .				
ASAQLENMEEAPKRVSLALQLPEHGSKDIĞN VPGNCSENPCQNGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV 1300 2650 A 10162 98 391 AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV 1301 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWI.GKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNFSKIEREHKRRTSTPV	1299	2649	Α	10161	1	393	PRFSELVDGRGRVSARFGGSPSKAATVRSOPT
1300 2650 A 10162 98 391 AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV 1301 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWI.GKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNFSKIEREHKRRTSTPV	•						ASAQLENMEEAPKRVSLALQLPEHGSKDIGN
KGRRCELACIKVSRPCTRLFSETKAFPVWEGG	[VPGNCSENPCQNGGTCVPGADAHSCDCGPGF
1300 2650 A 10162 98 391 AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV 1301 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWI.GKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGGPGSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNFSKIEREHKRRTSTPV	1		ĺ	1 1		}	KGRRCELACIKVSRPCTRLFSETKAFPVWEGG
1301 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWI.GKPDPGRKRRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNFSKIEREHKRRTSTPV							VCHHV
VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV 1301 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWI.GKPDPGRKRRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV	1300	2650	Α	10162	98	391	AKIASLERIMPANYTCTRPDGDNTDFRYFIYA
VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV 1301 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWI.GKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV							VTYTGILGPGLIGNILALWVFYGYMKETKRA
1301 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWI.GKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV							VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF
KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWI.GKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV							VPV
KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWI.GKPDPGRKRRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNFSKIEREHKRRTSTPV	1301	2651	A	10165	1	7545	PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE
EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWI.GKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNFSKIEREHKRRTSTPV						ľ	KQLQCMPMEGRGRASSSISDLQGKGFEKGTG
RWI.GKPDPGRKRRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNFSKIEREHKRRTSTPV						1	EKHVPGVGSARHSPQASAGGSPWQRGKAQT
LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV	j j				1	l	RWLGKPDPGRKRRRGSPQEEGGLRVSAAAR
PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV				' I	- 1	1	LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP
LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV					İ	1	PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA
RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV					1		LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP
DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV	i i	l		ł	ì		RTLSVEEPGVECNQLCLYADVTDPVLCLGOK
ARRLSESLHVVDENKNESKIEREHKRRTSTPV							DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP
IMEGVOEETDTRDVKROVER SHICTEEPOKO		j	ſ	1	ĺ	ĺ	ARRLSESLHVVDENKNESKIEREHKRRTSTPV
Tand (Qualitation) The Control of the Control of							IMEGVQEETDTRDVKRQVERSEICTEEPQKQ

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	ł	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K≈Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
]			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1 :				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
1 !			1			KSTLKNEKHLKKDDSETPHLKSLLKKEVKSS
1						KEKPEREKTPSEDKLSVKHKYKGDCMHKTG
1 1			i			DETELHSSEKGLKVEENIQKQSQQTKLSSDDK
						TERKSKHRNERKLSVLGKDGKPVSEYIIKTDE
ì i						NVRKENNKKERRLSAEKTKAEHKSRRSSDSK
1						IQKDSLGSKQHGITLQRRSESYSEDKCDMDST
1						NMDSNLKPEEVVHKEKRRTKSLLEEKLVLKS
1						KSKTQGKQVKVVETELQEGATKQATTPKPD KEKNTEENDSEKQRKSKVEDKPFEETGVEPV
						LETASSSAHSTQKDSSHRAKLPLAKEKYKSD
1 1						KDSTSTRLERKLSDGHKSRSLKHSSKDIKKKD
						ENKSDDKDGKEVDSSHEKARGNSSLMEKKL
]]						SRRLCENRRGSLSQEMAKGEEKLAANTLSTP
1						SGSSLQRPKKSGDMTLIPEQEPMEIDSEPGVE
1						NVFEVSKTQDNRNNNSHQDIDSENMKQKTS
i i						ATVOKDELRTCTADSKATAPAYKPGRGTGV
1						NSNSEKHADHRSTLTKKMHIQSAVSKMNPGE
1 1		l	1			KEPIHRGTTEVNIDSETVHRMLLSAPSENDRV
						QKNLKNTAAEEHVAQGDATLEHSTNLDSSPS
1 1			1			LSSVTVVPLRESYDPDVIPLFDKRTVLEGSTA
		ļ				STSPADHSALPNQSLTVRESEVLKTSDSKEGG
1 1					ì	EGFTVDTPAKASITSKRHIPEAHQATLLDGKQ
					ı	GKVIMPLGSKLTGVIVENENITKEGGLVDMA
1 1		-	.]		j	KKENDLNAEPNLKQTIKATVENGKKDGIAVD
))		1				HVVGLNTEKYAETVKLKHKRSPGKVKDISID
1			1		·	VERRNENSEVDTSAGSGSAPSVLHQRNGQTE DVATGPRRAEKTSVATSTEGKDKDVTLSPVK
		}	,	ļ		AGPATTTSSETRQSEVALPCTSIEADEGLIIGT
			ļ			HSRNNPLHVGAEASECTVFAAAEEGGAVVTE
			j			GFAESETFLTSTKEGESGECAVAESEDRAADL
			ſ			LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE
			1			KCDGSLSRDSEIVEGTITFISEVESDGAVTSAG
	Ī	ĺ	1	ĺ	•	TEIRAGSISSEEVDGSQGNMMRMGPKKETEG
						TVTCTGAEGRSDNFVICSVTGAGPREERMVT
1 1	1	1	ì	i	1	GAGVVLGDNDAPPGTSASQEGDGSVNDGTE
		İ				GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS
1		ł	-	1	j	SESEENGESAMDSTVAKEGTNVPLVAAGPCD
	ĺ					DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH
1	1	1	l	ļ	}	ASTCTGLGEESEGVLICESAEGDSQIGTVVEH
	1			. 1		VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE
	ļ	j		1		GPMTSAASDQSDSQLEKVEDTTISTGLVGGS
		ł	1	1	ł	YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE
Į l	1	Ī	1	j	Į	NEECDGLMATTASGDITNONSLAGGKNOGK
1	}	ł	ļ	1	ĺ	VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE
			1	ļ		ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS
		j	- 1]		RSEEKDECAMISTSIGEEFELPISSATTIKCAES
			1	!	ì	LQPVAAAVEERATGPVLISTADFEGPMPSAPP
]	J]	J	}	J	EAESPLASTSKEEKDECALISTSIAEECEASVS
				ŀ	, 1	GVVVESENERAGTVMEEKDGSGIISTSSVEDC
		{		İ	·	EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS
(1	ľ	ĺ	1	TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA
					}	Aliststaecmpisasidrheenqltadnpegn
		. !	ľ	i	i	GOLSATEVSKHKVPMPSLIAENNCRCPGPVR
			Į		ļ	GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL
		1	İ	Į.	ŀ	HLINAEEKNVLLNSLQKEDKSPETGTAGGSST
	İ	ì	ı	ŀ	ľ	ASYSAGRGLEGNANSPAHLRGPEQTSGOTAK
						DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA
}	1	1		ł		EHSFLPAEQQGSEDNLKTSTTKCITGQESKIAP
				l		, , , , , , , , , , , , , , , , , , ,

SEO ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid seguence (AssAlania C. C
NO: of	NO: of	hod	ID NO:	beginning		Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	1 1100	in NO:	nucleotide	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
	{ • •	J		location	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	l	USSN	10	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	ĺ	!	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	Į.	1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	ļ]	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1		1		peptide		/=possible nucleotide deletion, \=possible
L	1	l		sequence		nucleotide insertion
					<u> </u>	SHTMIPPATYSVALLAPKCEQDLTIKNDYSGK
1	1	1	1			WTDQASAEKTGDDNSTRKSFPEEGDIMVTVS
1	1	l	l	l .		SEENVCDIGNEESPLNVLGGLKLKANLKMEA
1				1		YVPSEEEKNGEILAPPESLCGGKPSGIAELQRE
J	j	1	İ	1		PLLVNESLNVENSGFRTNEEIHSEŞYNKGEISS
			1		1	GRKDNAEAISGHSVEADPKEVEEEERHMPKR
			ĺ	1		KDKURA GEDEBUDYBURA DEBREA COCO
	1	ļ	1	ł	,	KRKQHYLSSEDEPDDNPDVLDSRIETAQRQC
Ì			1	1		PETEPHATKEENSRDLEELPKTSSETNSTTSRV
1302	2652		10167	701	0.40	MEEKDEYSSSETTGEKPEQNDDDTIKSQE
1302	2652	Α	10167	321	842	EPSLFPFLRPSPARPPPRPPAPFPSPELAGPEPH
1			1			FVFYFFLSYVHPPKELAKYEYMEEQVILTEKG
1			İ	1		NSTVAGRGTSVRCLSPSPRPLPPLLPLLADLLE
([1			DGFGEHPFYHCLVAEVPKEHWTPEGNPSPFP
			1			EARETKCYVRSSVGCVEPLTTQAEVTENLDR
<u></u>			L	l		KNSQQVFKLLKKK
1303	2653	Α	10171	206	429	NMILLKKRRLLINSLGEGTINGLLDELLETNV
1			1			LSQEDTEIVKCENVTVIDKARDLLDSVIRKGA
			{	[RACEICITYI
1304	2654	A	10184	970	1524	LCTLSPGISGTAGSCLTTEPGTELGTSFAQNGF
				• • •	-5	YHEAVVLFTQALKLNPQDHRLFGNRSFCHER
						LGQPAWALADAQVALTLRPGWPRGLFRLGK
						AI MCI OPEDE A AAMOOFTI BOOGODDA ADDA
			1			ALMGLQRFREAAAVFQETLRGGSQPDAAREL
			ĺ]		RSCLLHLTLQGQRGGICAPPLSPGALQPLPHA
1305	2655		10104	2	204	ELAPSGLPSLRCPRSTALRSPGLSPLLH
1303	2000	Α	10194	2	394	TDLLGRRFRVDGAAMAACEGRRSGALGSSQ
]]		SDFLTPPVGGAPWAVATTVVMYPPPPPPPHR
]	ļ	DFISVTLSFGESYDNSKSWRRRSCWRKWKQL
					.	SRLQRNMILFLLAFLLFCGLLFYINLADHWKG
						IRNTCT
1306	2656	Α	10195	1	410	IPGSTISLEGPLSKWTNVMKGWQYRWFVLDY
						NAGLESYYTSKDKMMRGSRRGCVRLRGAVI
				İ		GIDDEDDSTFTITVDQKTFHFQARDADEREK
						WIHALEETILRHTLQLQVRVFTWFPDSSLVGA
						FFFWLVSGFFFK
1307	2657	A	10205	85	308	QGLPSTMVKLGCSFSGKPGKDPGDQDGAAM
j i				·	- 	DSVPLISPLDISQLQPPLPDQVVIKTQTEYQLS
						SPDQONYTKSR
1308	2658		10214	2	453	
ا "" ا	-050	••	10214	-	433 ,	ECGGIRQPGPPPALASAPAATMNRVGGSPS
! !					,	AAANYLLCTNCRKVLRKDKRIRVSQPLTRGP
					ŧ .	SAFIPEKEVVQANTVDERTNFLVEEYSTSGRL
						DNITQVMSLHTQYLESFLRSQFYMLRMDGPL
1309	260		10000	15		PLPYRHYIAIMAAARHQCSYLINM
1209	2659	Α	10233	45	421	RGWPEQQSTGRPRDVARQPRCQKEEGRRLRP
, ,	ļ	j			J	RALESRTFQGSERSRWGPPLESTKENVQCGH
<u> </u>						RPAFPNSSWLPFHERLQVQNGECPWQVSIQM
[.					j	SRKHLCGGSILHWWWVLTAAHCFRRTLLDM
					*	AV
1310	2660	A	10241	243	442	AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK
	ŀ	ł	· '']	ļ		HKKGQSAEIQKKRTRRAFKFQRAITGASLADI
	}			ì		MAK
1311	2661	Ā	10261	751	176	LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ
(- 1				VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ
	Ì	1				LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF
	ł				1	OVDVIDIONICOLOGICALINATIVO CONTRACTOR INCOME
	ļ	ł	1	ļ		QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI
	İ	ŀ		l	}	DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML
	ĺ	ĺ	ĺ	!	ľ	QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI
1210	3662		10000			KAG
1312	2662	A	10270	3	669	STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP
						SMTILDKKDGEQAKALFEKVRKFRAHVEDSD

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LIYKLYVVQTVIKTAKFIFILCYTANFVNAISF EHVCKPKVEHLIGYEVFECTHNMAYMLKKL LISYISIICVYGFICLYTLFWLFRIPLKEYSFEKV REESSFSDIPDVKNDFAFLLHMVDQYDQLYS KRFGVFLSEVSENKLREISLNHEWTFEKL
1313	2663	A	10287	1221	266	GAHRVLSPAQGAQPRLRSAASVEVSMVGQR VLLLVAFLLSGVLLSEAAKILTISTLGGSHYLL LDRVSQILQEHGHNVTMLHQSGKFLIPDIKEE EKSYQVIRWFSPEDHQKRIKKHFDSYIETALD GRKESEALVKLMEIFGTQCSYLLSRKDIMDSL KNENYDLVFVEAFDFCSFLIAEKLVKPFVAIL PTTFGSLDFGLPSPLSYVPVFPSLLTDHMDFW GRVKNFLMFFSFSRSQWDMQSTFDNTIKEHF PEGSRPVLSHLLLKAELWFVNSDCAFDFARPL LPNTVYIGGLMEKPIKPVPQVSEPSAFSLGFT
1314	2664	A	10288	536	1890	NVQLAKFSSTLVFFFSCDADPSALAKYVLAL VKKDKSEKELKALCIDQLDVFLQKETQIFVEK LFDAVNTKSYLPPPEQPSSGSLK VEFFPPQEK DIKKEEITKEEEREKKFSRRLNHSPPQSSSRYR ENRSRDERKKDDRSRKRDYDRNPPRRDSYRD RYNRRRGRSRSYSRSRSRSWSKERLRERDRD RSRTRSRSRTRSRERDLVKPKYDLDRTDPLEN NYTPVSSVPSISSGHYPVPTLSSTITVIAPTHHG NNTTESWSEFHEDQVDHNSYVRPPMPKKRC RDYDEKGFCMRGDMCPFDHGSDPVVVEDVN LPGMQPFPAQPPVVEGPPPPGLPPPPILTPPPV NLRPPVPPPGPLPPSG MDAPPNSATSSVPTVVTTGIHHQPPPAPPSLFT ADTYDTDGYNPEAPSITNTSRPMYRHRVHPR AKLG
1315	2665	A	10293	447	1331	SHPLLSCPEKVSAKLRAAAEAAAERRTRGA GSRGICAGLRSVAPGPEPLKQEEGRREWGSSI GTPSPCGSAQAAAAAAAEEATEKIPALRPALL WALLALWLCCATPAHALQCRDGYEPCVNEG MCVTYHNGTGYCKCPEGFLGEYCQHRDPCE KNRCQNGGTCVAQAMLGKATCRCASGFTGE DCQYSTSHPCFVSRPCLNGGTCHMLSRDTYE CTCQVGFTGRNPKCPGGNLNYQFNGIIVYYS GGSVPPSGTKTSKPAEHNAMGTGSKNFASGT LWVMVSGATSTSTSTL
1316	2666	A	10294	118	572	SLSMESNHKSGDGLSGTQKEAALRALVQRTG YSLVQENGQRKYGGPPPGWDAAPPERGCEIFI GKLPRDLFEDELIPLCEKIGKIYEMRMMMDF NGNNRGYAFVTFSNKVEAKNAIKQLNNYEIR NGRLLGVCASVDNCRLFVGGIPKTKK
1317	2667	A	10301	158	1956	LLKSCGVLLSGVCIPCEGKGPTVLVIQTAVPQ DRPTKSSMRSAAKPWNPAIRAGGHGPDRVRP LPAASSGMKSSKSSTSLAFESRLSRLKRASSE DTLNKPGSTAASGVVRLKKTATAGAISELTES RLRSGTGAFTTTKRTGIPAPREFSVTVSRERSV PRGPSNPRKSVSSPTSSNTPTPTKHLRTPSTKP KQENEGGEKAALESQVRELLAEAKAKDSEIN RLRSELKKYKEKRTLNAEGTDALGPNVDGTS VSPGDTEPMIRALEEKNKNFQKELSDLEEENR VLKEKLIYLEHSPNSEGAASHTGDSSCPTSITQ ESSFGSPTGNQLSSDIDEYKKNIHGNALRTSG SSSSDVTKASLSPDASDFEHITAETPSRPLSSTS NPFKSSKCSTAGSSPNSVSELSLASLTEKIQKM EENHHSTAEELQATLQELSDQQQMVQELTAE

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine.
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	İ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ļ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
			1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon.
ł		1		peptide	Soquence	/=possible nucleotide deletion, \=possible
1				sequence	•	nucleotide insertion
ļ	 		 	sequence	 	NEKLVDEKTILETSFHOHRERAEOLSOENEKL
İ		J			ļ	
			 			MNLLQERVKNEEPTTQEGKIIELEQKCTGILE
					i	QGRFEREKLLNIQQQLTCSLRKVEEENQGAL
1		1	ł			EMIKRLKEENEKLNEFLELERHNNNMMAKTL
1210	2220		10000			EECRVTLEGLKMENGSLKSHLQG
1318	2668	Α	10303	333	879	GECFIMAAVVQQNDLVFEFASNVMEDERQL
			1			GDPAIFPAVIVEHVPGADILNSYAGLACVEEP
						NDMITESSLDVAEEEIIDDDDDDITLTVEASCH
1						DGDETIETIEAAEALLNMDSPGPMLDEKRINN
		1	1			NIFSSPEDDMVVAPVTHVSVTLDGIPEVMETQ
			<u> </u>			QVQEKYADSPGASSPEQPKRKKK
1319	2669	A	10322	169	654	MEVRMSGSVAVTRAIAVPGLLLLIIATALSL
i						LIGAKSLPASVVLEAFSGTCQSADCTIVLDAR
						LPRTLAGLLAGGALGLAGALMQTLTRNPLAD
		[[(PGLLGVNAGASFAIVLGAALFGYSSAQEQLA
		1	ļ			MAFAGALVASLIVAFTGSQGGGQLSPVRLTL
1						AGVXL
1320	2670	A	10323	441	2	KMNQVAVVIGGGQTLGAFLCHGLAAEGYRV
1	20.0	••	10323		-	AVVDIQSDKAANVAQEINAEYGESMAYGFG
	•	1	ĺ			ADATSEQSVLALSRGVDEIFGRVDLLVYSAGI
						AKAAFISDFQLGDFDRSLQVNLVGYFLCARE
1	i	ł				FSRLMIRDGIQGRIIQINSKSDE
1321	2671	A	10332	1	453	
1321	2071	^	10332	1	453	RHRTAGPGSTISSRTDSASAPAARAMPCEYTY
]				AKLTSDCSRPSLQWYTRAQSKMRRPRLLLKD
						ILKCTLLVFGVRILYILKLNYTTEECDMKNMH
		ļ				YVDPDHVKRAQKYAQQVLQKESPPKFAKTS
1322	0.00	-	10222			MALLFEHRYSVDLLPFVQKAPTDSEA
1322	2672	Α	10333	25	423	EPSNGPVVYSALGNEDDEILLLGKDIIGTFAAS
1		1				ERKMRAHQVLTFLLLFVITSGASENASTSRGC
 			"			GLDLLPQNVYLCDLDAIWGIVVEAVAGAGA
		}				<u>LITLLMLILLGRLPFIKEKEKKSPAVLHFLFL</u>
						LGTLG
1323	2673	Α	10334	52	426	SSLGNEDDEILSLAKDITGMFVASHRKMRAH
1						QVLTFLLLFVITSVASENASTSRGCGLDLLPQ
						YVSLCDLDAIWGIVVEAAAGAGALITLLLMLI
1.						LLVRLPFFKEKEKKSPVGLHFLFLLGTLGP
1324	2674	Α	10336	1	932	ERLCFPCMQSKIYSYMSPNKCSGMRFPLQEE
1 1		,				NSVTHHEVKCOGKPLAGIYRKREEKRNAGN
1						AVRSAMKSEEQKIKDARKGPLVPFPNQKSEA
Į l						AEPPKTPPSSCDSTNAAIAKQALKKPIKGKQA
						PRKKAQGKTQQNRKLTDFYPVRRSSRKSKAE
1 !						LQSEERKRIDELIESGKEEGMKIDLIDGKGRG
	,					VIATKQFSRGDFVVEYHGDLIEITDAKKREAL
					İ	YAQDPSTGCYMYYFQYLSKTYCVDATRETN
1 1						RLGRLINHSKCGNCQTKLHDIDGVPHLILIAS
1325	2675	Α	10338	3	970	RDIAAGEELLYDYGDRSKASIEAHPWLKH
(26.1	2013	^	Ιυσοδ	د	870	PGSTISCSELKGTQCRATAGSRGRRPPMTCWL
						RGVTATFGRPAEWPGYLSHLCGRSAAMDLG
						PMRKSYRGDREAFEETHLTSLDPVKQFAAWF
1 1			1			EEAVQCPDIGEANAMCLATCTRDGKPSARML
						LLKGFGKDGFRFFTNFESRKGKELDSNPFASL
			ı j	ļ	ļ	VFYWEPLNRQVRVEGPVKKLPEEEAECYFHS
1 !			.]	ļ	ļ	RPKSSQIGAVVSHQSSVIPDREYLRKKNEELE
	.			l		QLYQDQEVPKPKSWGGYVLYPQVMEFWQG
1 1				Į	i	QTNRLHDRIVFRRGLPTGDSPLGPMTHRGEE
i		l	Ì	l	ľ	DWLYERLAP
1326	2676	Α	10344	2	984	ARAAAHCGICRLVRWWRKRRSVMGIQTSPV
1 1			1	}		LLASLGVGLVTLLGLAVGSYLVRRSRRPQVT
			i			LLDPNEKYLLRLLDKTTVSHNTKRFRFALPTA

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion HHTLGLPVGKHIYLSTRIDGSLVIRPYTPVTSD EDQGYVDLVIKVYLKGVHPKFPEGGKMSQY LDSLKVGDVVEFRGPSGLLTYTGKGHFNIQP
1327	2677	A	10345	1	968	NKKSPPEPRVAKKLGMIAGGTGITPMLQLIRA ILKVPEDPTQCFLLFANQTEKDIILREDLEELQ ARYPNRFKLWFTLDHPPKDWAYSKGFVTAD MIREHLPAPGDDVLVLLCGPPPMVQLACHPN LDKLGYSQKMRFTY LQSAGEGVTHVLILLESPARPVAAVTQVQRR
						RYHRLSDMSMLAERRRKQKWAVDPQNTAW SNDDSKFGQRMLEKMGWSKGKGLGAQEQG ATDHIKVQVKNNHLGLGATINNEDNWIAHQ DDFNQLLAELNTCHGQETTDSSDKKEKKSFS LEEKSKISKNRVHYMKFTKGKDLSSRSKTDL DCIFGKRQSKKTPEGDASPSTPEENETTTTSAF TIQEYFAKRMAALKNKPQVPVPGSDISETQVE RKRGKKRNKEATGKDVESYLQPKAKRHTEG KPERAEAQERVAKKKSAPAEEQLRGPCWDQ SSKASAQDAGDHVQPA
1328	2678	Α	10346	173	439	GSAAMKVKIKCWNGVATWLWVANDENCGI CRMAFNGCCPDCKVPGDDCPLVWGQCSHCF HMHCILKWLHAQQVQQHCPMCRQEWKFKE
1329	2679	A	10351	3	964	QMEPGNDTQISEFLLLGFSQEPGLQPFLFGLFL SMYLVTVLGNLLIILATISDSHLHTPMYFFLSN LSFADICVTSTTIPKMLMNIQTQNKVTTYIACL MQMYFFILFAGFENFLLSVMAYDRFVAICHP LHYMVIMNPHLCGLLVLASWTMSALYSLLQI LMVVRLSFCTALEIPHFFCELNQVIQLACSDSF LNHMVIYFTVALLGGGPLTGILYSYSKIISSIH AISSAQGKYKAFSTCASHLSVVSLFYGAILGV YLSSAATRNSHSSATASVMYTVVTPMLNPFI YSLRNKDIKRALGIHLLWGTMKGQFFKKCP
1330	2680	A	10352	34	2573	IPFLKSCCCCCLFDFPPPPLDQVQEEECEVERV TEHGTPKPFRKFDSVAFGESQSEDEQFENDLE TDPPNWQQLVSREVLLGLKPCEIKRQEVINEL FYTERAHVRTLKVLDQVFYQRVSREGILSPSE LRKIFSNLEDILQLHIGLNEQMKAVRKRNETS VIDQIGEDLLTWFSGPGEEKLKHAAATFCSNQ PFALEMIKSRQKKDSRFQTFVQDAESNPLCRR LQLKDIIPTQMQRLTKYPLLLDNIATYTEWPT EREKVKKAADHCRQILNYVNQAVKEAENKQ RLEDYQRRLDTSSLKLSEYPNVEELRNLDLTK RKMIHEGPLVWKVNRDKTIDLYTLLLEDILV LLQKQDDRLVLRCHSKILASTADSKHTFSPVI KLSTVLVRQVATDNKALFVISMSDNGAQIYE LVAQTVSEKTVWQDLICRMAASVKEQSTKPI PLPQSTPGEGDNDEEDPSKLKEEQHGISVTGL QSPDRDLGLESTLISSKPQSHSLSTSGKSEVRD LFVAERQFAKEQHTDGTLKEVGEDYQIAIPDS HI.PVSEERWALDALRNLGLLKQLI.VQQLGLT EKSVQEDWQHFPRYRTASQGPQTDSVIQNSE NIKAYHSGEGHMPFRTGTGDIATCYSPRTSTE SFAPRDSVGLAPQDSQASNILVMDHMIMTPE MPTMEPEGGLDDSGEHFFDAREAHSDENPSE GDGAVNKEEKDVNLRISGNYLILDGYDPVQE SSTDEEVASSLTLQPMTGIPAVESTHQQHSP QNTHSDGAISPFTPEFLVQQR WGAMEYSCFEI QSPSSCADSQSQIMEYIHKIEADLEHLKKVEE

SEQ ID NO: of nucl-otide peptide eotide sequence wence	ine, ne, n, in, ine DKYLEG VTKIELLP SGIKWSE FVCSLDIL LLGLVIG TVRAAIP FFRAFA YLEIITQL LIVQLDK IFTNKTQ TMKNVT LLILSLLV
nucleotide sequence in nucleotide location corresponding to last amino acid residue of peptide sequence in sequence in sequence in corresponding to last amino acid residue of peptide sequence in seq	ne, n, n, ne DKYLEG VTKIELLP SGIKWSE FVCSLDIL PLIGLVIG TVRAAIP FRRAFA YLEIITQL LLIVQLDK IFTNKTQ TMKNVT LLILSLLV
eotide sequence Sequence USSN 09/496 914 14 16 16 18 18 18 18 18 18	ne, n, n, ne DKYLEG VTKIELLP SGIKWSE FVCSLDIL PLIGLVIG TVRAAIP FRRAFA YLEIITQL LLIVQLDK IFTNKTQ TMKNVT LLILSLLV
sequence Sequence 09/496	DKYLEG OKYLEG VTKIELLP SGIKWSE FVCSLDIL PLIGLVIG TVRAAIP FRRAFA YLEIITQL LLIVQLDK IFTNKTQ TMKNVT LLILSLLV
ng to first amino acid residue of peptide sequence 1331 2681 A 10353 I 2100 AVEFAEGALTMAPWPELGDAQPNP AAGQQPTAPDKSKETNKTDNTEAP SYSTATLIDEPTEVDDPWNLPTLQD RDTKGKILCFFQGIGRLILLGFLYF SSAFQLVGGKMAGQFSNSSIMSNP VLVTVLVQSSSTSTSIVVSMVSSSLL IIMGANIGTSITNTIVALMQVGDRSE GATVHDFFNWLSVLVLLPVEVATH IVESFHFKNGEDAPDLLKVITKPFTK KVISQIAMNDEKAKNKSJ VKIWCK INVTVPSTANCTSPSLCWTDGIQNW YKENIAKCQHIFVNFHLPDLAVGTII LCGCLIMIVKILGSVLKGQVATVIKK FPFAWLTGYLAILVGAGMTFIVQSS	DKYLEG OKYLEG VTKIELLP SGIKWSE FVCSLDIL PLIGLVIG TVRAAIP FRRAFA YLEIITQL LLIVQLDK IFTNKTQ TMKNVT LLILSLLV
amino acid residue of peptide sequence	DKYLEG VTKIELLP SGIKWSE FVCSLDIL PLLGLVIG TVRAAIP FRRAFA YLEIITQL LLIVQLDK IFTNKTQ TMKNVT LLILSLLV
residue of peptide sequence Y=Tyrosine, X=Unknown, *=Stop codo /=possible nucleotide deletion, \=possible nucleotide insertion 1331 2681 A 10353 I 2100 AVEFAEGALTMAPWPELGDAQPNP AAGQQPTAPDKSKETNKTDNTEAP SYSTATLIDEPTEVDDPWNLPTLQD: RDTKGKILCFFQGIGRLILLLGFLYFI SSAFQLVGGKMAGQFFSNSSIMSNP VLVTVLVQSSSTSTSIVVSMVSSSLL IIMGANIGTSITNTIVALMQVGDRSE GATVHDFFNWLSVLVLLPVEVATH IVESFHFKNGEDAPDLLKVITKPFTK KVISQIAMNDEKAKNKSJVKIWCKI INVTVPSTANCTSPSLCWTDGIQNW YKENIAKCQHIFVNFHLPDLAVGTII LCGCLIMIVKILGSVLKGQVATVIKK FPFAWLTGYLAILVGAGMTFIVQSS	DKYLEG VTKIELLP SGIKWSE FVCSLDIL PLLGLVIG TVRAAIP FRRAFA YLEIITQL LLIVQLDK IFTNKTQ TMKNVT LLILSLLV
peptide /=possible nucleotide deletion, \=possible nucleotide insertion	DKYLEG VTKIELLP SGIK WSE FVCSLDIL PLLGLVIG TVRAAIP FRRAFA YLEIITQL LLIVQLDK IFTNKTQ TMKNVT LLILSLLV
1331 2681 A 10353 I 2100 AVEFAEGALTMAPWPELGDAQPNP AAGQQPTAPDKSKETNKTDNTEAP SYSTATLIDEPTEVDDPWNLPTLQD RDTKGKILCFFQGIGRLILLLGFLYFF SSAFQLVGGKMAGQFFSNSSIMSNP VLVTVLVQSSSTSTSIVVSMVSSSLL IIMGANIGTSITNTIVALMQVGDRSE GATVHDFFNWLSVLVLLPVEVATH IVESFHFKNGEDAPDLLKVITKPFTK KVISQIAMNDEKAKNKSLVKIWCKT INVTVPSTANCTSPSLCWTDGIQNW YKENIAKCQHIFVNFHLPDLAVGTII LCGCLIMIVKILGSVLKGQVATVIKK FPFAWLTGYLAILVGAGMTFIVQSS	PDKYLEG VTKIELLP SGIKWSE FVCSLDIL PLEGLVIG TVRAAIP FRRAFA YLEIITQL LIVQLDK IFTNKTQ TMKNVT LILSLLV
AAGQQPTAPDKSKETNKTDNTEAP' SYSTATLIDEPTEVDDPWNLPTLQD: RDTKGKILCFFQGIGRLILLLGFLYFI SSAFQLVGGKMAGQFFSNSSIMSNP VLVTVLVQSSSTSTSIVVSMVSSSLL IIMGANIGTSITNTIVALMQVGDRSE GATVHDFFNWLSVLVLLPVEVATH IVESFHFKNGEDAPDLLKVITKPFTK KVISQIAMNDEKAKNKSLVKIWCKI INVTVPSTANCTSPSLCWTDGIQNW YKENIAKCQHIFVNFHLPDLAVGTII LCGCLIMIVKILGSVLKGQVATVIKK FPFAWLTGYLAILVGAGMTFIVQSS	VTKIELLP SGIKWSE FVCSLDIL PLIGLVIG TVRAAIP FRRAFA YLEIITQL LIIVQLDK IFTNKTQ TMKNVT LILSLLV
AAGQQPTAPDKSKETNKTDNTEAP' SYSTATLIDEPTEVDDPWNLPTLQD: RDTKGKILCFFQGIGRLILLLGFLYFI SSAFQLVGGKMAGQFFSNSSIMSNP VLVTVLVQSSSTSTSIVVSMVSSSLL IIMGANIGTSITNTIVALMQVGDRSE GATVHDFFNWLSVLVLLPVEVATH IVESFHFKNGEDAPDLLKVITKPFTK KVISQIAMNDEKAKNKSLVKIWCKT INVTVPSTANCTSPSLCWTDGIQNW YKENIAKCQHIFVNFHLPDLAVGTII LCGCLIMIVKILGSVLKGQVATVIKK FPFAWLTGYLAILVGAGMTFIVQSS	VTKIELLP SGIKWSE FVCSLDIL PLLGLVIG TVRAAIP FRRAFA YLEIITQL LLIVQLDK IFTNKTQ TMKNVT LLILSLLV
RDTKGKILCFFQGIGRLILLLGFLYFI SSAFQLVGGKMAGQFFSNSSIMSNP VLVTVLVQSSSTSTSIVVSMVSSSLL IIMGANIGTSITNTIVALMQVGDRSE GATVHDFFNWLSVLVLLPVEVATH IVESFHFKNGEDAPDLLKVITKPFTK KVISQIAMNDEKAKNKSLVKIWCKT INVTVPSTANCTSPSLCWTDGIQNW YKENIAKCQHIFVNFHLPDLAVGTII LCGCLIMIVKILGSVLKGQVATVIKK FPFAWLTGYLAILVGAGMTFIVQSS	FVCSLDIL PLEGLVIG TVRAAIP FRRAFA YLEIITQL LIVQLDK IFTNKTQ TMKNVT LILSLLV
SSAFQLVGGKMAGQFFSNSSIMSNP VLVTVLVQSSSTSTSIVVSMVSSSLL IIMGANIGTSITNTIVALMQVGDRSE GATVHDFFNWLSVLVLLPVEVATH IVESFHFKNGEDAPDLLKVITKPFTK KVISQIAMNDEKAKNKSLVKIWCKT INVTVPSTANCTSPSLCWTDGIQNW YKENIAKCQHIFVNFHLPDLAVGTII LCGCLIMIVKILGSVLKGQVATVIKK FPFAWLTGYLAILVGAGMTFIVQSS	LLGLVIG TVRAAIP FRRAFA YLEIITQL LIVQLDK IFTNKTQ TMKNVT LLILSLLV
VLVTVLVQSSSTSTSIVVSMVSSSLL IIMGANIGTSITNTIVALMQVGDRSE GATVHDFFNWLSVLVLLPVEVATH IVESFHFKNGEDAPDLLKVITKPFTK KVISQIAMNDEKAKNKSLVKIWCKI INVTVPSTANCTSPSLCWTDGIQNW YKENIAKCQHIFVNFHLPDLAVGTII LCGCLIMIVKILGSVLKGQVATVIKK FPFAWLTGYLAILVGAGMTFIVQSS	TVRAAIP FRRAFA YLEIITQL LIVQLDK FFTNKTQ TMKNVT LILSLLV
IIMGANIGTSITNTIVALMQVGDRSE GATVHDFFNWLSVLVLLPVEVATH IVESFHFKNGEDAPDLLKVITKPFTK KVISQIAMNDEKAKNKSLVKIWCKT INVTVPSTANCTSPSLCWTDGIQNW YKENIAKCQHIFVNFHLPDLAVGTII LCGCLIMIVKILGSVLKGQVATVIKK FPFAWLTGYLAILVGAGMTFIVQSS	FRRAFA YLEIITQL LIVQLDK FTNKTQ TMKNVT LLILSLLV
GATVHDFFNWLSVLVLLPVEVATH IVESFHFKNGEDAPDLLKVITKPFTK KVISQIAMNDEKAKNKSLVKIWCKT INVTVPSTANCTSPSLCWTDGIQNW YKENIAKCQHIFVNFHLPDLAVGTII LCGCLIMIVKILGSVLKGQVATVIKK FPFAWLTGYLAILVGAGMTFIVQSS	YLEIITQL LIVQLDK IFTNKTQ TMKNVT LIILSLLV
IVESFHFKNGEDAPDLLKVITKPFTK KVISQIAMNDEKAKNKSLVKIWCKT INVTVPSTANCTSPSLCWTDGIQNW YKENIAKCQHIFVNFHLPDLAVGTII LCGCLIMIVKILGSVLKGQVATVIKK FPFAWLTGYLAILVGAGMTFIVQSS	LIVQLDK FFTNKTQ TMKNVT LLILSLLV
KVISQIAMNDEKAKNKSLVKIWCKT INVTVPSTANCTSPSLCWTDGIQNW YKENIAKCQHIFVNFHLPDLAVGTII LCGCLIMIVKILGSVLKGQVATVIKK FPFAWLTGYLAILVGAGMTFIVQSS	TFTNKTQ TMKNVT LLILSLLV
INVTVPSTANCTSPSLCWTDGIQNW YKENIAKCQHIFVNFHLPDLAVGTII LCGCLIMIVKILGSVLKGQVATVIKK FPFAWLTGYLAILVGAGMTFIVQSS	TMKNVT LLILSLLV
YKENIAKCQHIFVNFHLPDLAVGTII LCGCLIMIVKILGSVLKGQVATVIKK FPFAWLTGYLAILVGAGMTFIVQSS	LLILSLLV
LCGCLIMIVKILGSVLKGQVATVIKK FPFAWLTGYLAILVGAGMTFIVQSS	
FPFAWLTGYLAILVGAGMTFIVQSS	TINTDEP
TPLIGIGVITIERAYPLTLGSNIGTTT	SVFTSAL
	TAILAAL
ASPGNALRSSLQIALCHFFFNISGILL	WYPIPFT
RLPIRMAKGLGNISAKYRWFAVFYI	LIIFFFLIP
LTVFGLSLAGWRVLVGVGVPVVFII	ILVLCLR
LLQSRCPRVLPKKLQNWNFLPLWM	RSLKPW
DAVVSKFTGCFQMRCCCCCRVCCR	ACCLLC
GCPKCCRCSKCCEDLEEAQEGQDVI	
1332 2682 A 10354 30 1377 SOOGSOPHROGERS I TARRISTOLE	
1 SOUTH TO S	ALPPGPR
GSQGKLRRVLVPMSVKPSWGPGPSI	EGVTAVP
TSDLGEIHNWTELLDLFNHTLSECH	VELSQST
KRVVLFALYLAMFVVGLVENLLVIQ	JVNWRG
SGRAGLMNLYILNMAIADLGIVLSL	rvwMLE
VTLDYTWLWGSFSCRFTHYFYFVNI LVCLSVDRYVTLTSASPSWQRYQHI	MYSSLFF
CAGIWVLSAIIPLPEVVHIQLVEGPEI	
APFETYSTWALAVALSTTILGFLLPF	DITALN
VLTACRLRQPGQPKSRRHCLLLCAY	
MCWLPYHVTLLLLTLHGTHISLHCH	ILVHLLY
FFYDVIDCFSMLHCVINPILYNFLSPH	HFRGRLI.
NAVVHYLPKDQTKAGTCASSSSCST	THISHOT
KGDSQPAAAAPHPEPSLSFQAHHLL	
TQPLTPS	ľ
1333 2683 A 10358 2 884 AAGAGADGREPASERASRAEPPAVA	
LMGTAEDFADQFLRVTKQYLPHVA	RLCLIST
	CGYLLA
SSFVFLNLLGQLTGCVLVLSRNFVQ	YACFGLF
GIIALQTIAYSILWDLKFLMRNLALG	GGLLLL
LAESRSEGKSMFAGVPTMRESSPKQ	YMQLGG
RVLLVLMFMTLLHFDASFFSIVQNIV	GTALMI
LVAIGFKTKLAALTLVVWLFAINVY	FNAFWT
IPVYKPMHDFLKYDFFQTMSVIGGLI	LLVVAL
GPGGVSMDEKKKEW 1334 2684 A 10367 59 1562 OAWSLOVALSPFFPASPSNSFAAAV	(001177
1334 2684 A 10367 59 1562 QAWSLQVALSPFFFPASPSNSFAAAV ELPLPHVPGQESAKRRSARRFLIMSE	VEQLLEP
ELFLPHVFGQESAKRKSAKRFLIMSE	TIKETW
ALEQFEGGPCAVIAPVQAFLLKKLLF	COEDEO
WRDCSQEEQKELLCHTLCDILESACO	CDH6G6
YCL VS WLRGKTTEET AS IS GSPAESS	CONERG
SALAVEELGFERFHALIQKRSFRSLPI	ELKDAN
LDQYSMWGNKFGVLLFLYSVLLTKO	GIENIKN
EIEDASEPLIDPVYGHGSQSLINLLLT	GHAVSN
VWDGDRECSGMKLLGIHEQAAVGF	LTLMEA
LRYCKVGSYLKISKIPYLDCLASETH	
KDMALVAPEAPSEQARRVFQTYDPE	DNGFIP
DSLLEDVMKALDLVSDPEYINLMKN	KLDPEG

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, I.=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide deletion, \=DILLGPFLQEFFPDQGSSGPESFTVYHYNGL KQSNYNEKVMYVEGTAVVMGFEDPMLQTD DTPIKRCLQTKWPYIELLWTTDRSPSLN
1335	2685	A	10375	82	2929	TRTKRRLGREKAMASPPRGWGCGELLIPFML LGTLCEPGSGQIRYSMPEELDKGSFVGNIAKD LGLEPQELAERGVRIVSRGRTQLFALNPRSGS LVTAGRIDREELCAQSPLCVVNFNILVENKM KIYGVEVEIIDINDNFPRFRDEELKVKVNENA AAGTRLVLPFARDADVGVNSLRSYQLSSNLH FSLDVVSGTDGQKYPELVLEQPLDREKETVH DLLLTALDGGDPVLSGTTHIRVTVLDANDNA PLFTPSEYSVSVPENIPVGTRLLMLTATDPDE GINGKLTYSFRNEEEKISETFQLDSNLGEISTL QSLDYEESRFYLMEVVAQDGGALVASAKVV VTVQDVNDNAPEVILTSLTSSISEDCLPGTVIA LFSVHDGDSGENGEIACSIPRNLPFKLEKSVD NYYHLLTTRDLDREETSDYNITLTVMDHGTP PLSTESHIPLKVADVNDNPPNFPQASYSTSVT ENNPRGVSIFSVTAHDPDSGDNARVTYSLAE DTFQGAPLSSYVSINSDTGVLYALRSFDYEQL RDLQLWVTASDSGNPPLSSNVSLSLFVLDQN DNTPEILYPALPTDGSTGVELAPRSAEPGYLV TKVVAVDKDSGQNAWLSYRLLKASEPGLFA VGLHTGEVRTARALLDRDALKQSLVVAVED HGQPPLSATFTVTVAVADRIPDILADLGSIKTP IDPEDLDLTLYLVVAVAAVSCVFLAFVIVLLV LRLRRWHKSRLLQAEGSRLAGVPASHFVGV DGVRAFLQTYSHEVSLTADSRKSHLIFPQPNY ADTLLSESCEKSEPLLMSDKVDANKEERRV QQAPPNTDWRFSQAQRPGTSGSQNGDDTGT WPNNQFDTEMLQAMILASASEAADGSSTLGG GAGTMGLSARYGPQFTLQHVLQGELGSDYR QNVYIPGSNATLTNAAGKRDGKAPAGGNGN KKKSGKKEKK
1336	2686	A	10379	1	557	RPRRRQPSFSCRVLVLEDPPCFRFTNSMNQEK LAKLQAQVRIGGKGTARRKKKVVHRTATAD DKKLQSSLKKLAVNNIAGIEEVNMIKDDGTVI HFNNPKVQASLSANTFAITGHAEAKPITEMLP GILSQLGADSLTSLRKLAEQFPRQVLDSKAPK PEDIDEEDDDVPDLVENFDEASKNEAN
1337	2687	A	10380		1263	IPGSTISWSPAAARGLSVCRCCRLHPASAMDL FGDLPEPERSPRPAAGKEAQKGPLLFDDLPPA SSTDSGSGGPLLFDDLPPASSGDSGSLATSISQ MVKTEGKGAKRKTSEEEKNGSEELVEKKVC KASSVIFGLKGYVAERKGEREEMQDAHVILN DITECRPPSSLITRVSYFAVFDGHGGIRASKF AAQNLHQNLIRKFPKGDVISVEKTVKRCLLD TFKHTDEEFLKQASSQKPAWKDGSTATCVLA VDNILYIANLGDSRAILCRYNEESQKHAALSL SKEHNPTQYEERMRIQKAGGNVRDGRVLGV LEVSRSIGDGQYKRCGVTSVPDIRRCQLTPND RFILLACDGLFKVFTPEEAVNFILSCLEDEKIQ TREGKSAADARYEAACNRLANKAVQRGSAD NVTVMVVRIGH
1338	2688	A	10385	3	589	GPSQSMAAGELEGGKPLSGLLNALAQDTFHG YPGITEELLRSQLYPEVPPEEFRPFLAKMRGIL KSIASADMDFNQLEAFLTAQTKKQGGITSDQ AAVISKFWKSHKTKIRESLMNQSRWNSGLRG LSWRVDGKSQSRHSAQIHTPVAIIELELGKYG QESEFLCLEFDEVKVNQILKTLSEVEESISTLIS

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion QPN
1339	2689	A	10386	50	390	LGAMAKHHPDLIFCRKQAGVAIGRLCEKCDG KCVICDSYVRPCTLVRICDECNYGSYQGRCVI CGGPGVSDAYYCKECTIQEKDRDGCPKIVNL GSSKTDLFYERKKYGFKKR
1340	2690		10388	113	3472	SQLRKGASATHSSPSRTDCIAQMMDIYVCLK RPSWMVDNKRMRTASNFQWLLSTFILLYLM NQVNSQKKGAPHDLKCVTNNLQVWNCSWK APSGTGRGTDYEVCIENRSRSCYQLEKTSIKIP ALSHGDYEITINSLHDFGSSTSKFTLNEQNVSL IPDTPEILNLSADFSTSTLYLKWNDRGSVFPHR SNVIWEIKVLRKESMELVKLVTHNTTLNGKD TLHHWSWASDMPLECAIHFVEIRCYIDNLHFS GLEEWSDWSPVKNISWIPDSQTKVFPQDKVIL VGSDITFCCVSQEKVLSALIGHTNCPLIHLDGE NVAIKIRNISVSASSGTNVVFTTEDNIFGTVIF AGYPPDTPQQLNCETHDLKEIICSWNPGRVTA LVGPRATSYTLVESFSGKYVRLKRAEAPTNES YQLLFQMLPNQEIYNFTLNAHNPLGRSQSTIL VNITEKVYPHTPTSFKVKDINSTAVKLSWHLP GNFAKINFLCEIEIKKSNSVQEQRNVTIKGVE NSSYLVALDKLNPYTLYTFRIRCSTETFWKW SKWSNKKQHLTTEASPSKGPDTWREWSSDG KNLIIYWKPLPINEANGKILSYNVSCSSDEETQ SLSEIPDPQHKAEIRLDKNDYIISVVAKNSVGS SPPSKIASMEIPNDDLKIEQVVGMGKGILLTW HYDPNMTCDYVIKWCNSSRSEPCLMDWRKV PSNSTETVIESDEFRPGIRYNFFLYGCRNQGY QLLRSMIGYIEELAPIVAPNFTVEDTSADSILV KWEDIPVEELRGFLRGYLFYFGKGERDTSKM
			·		i ,	RVLESGRSDIKVKNITDISQKTLRIADLQGKTS YHLVLRAYTDGGVGPEKSMYVVTKENSVGL IIALLIPVAVAVIVGVVTSILCYRKREWIKETFY PDIPNPENCKALQFQKSVCEGSSALKTLEMNP CTPNNVEVLETRSAFPKIEDTEIVSPVAERPEN RSDAKPENHVVESYCPPIIEEEIPNPAADETGG TAQVIYIDVQSMYQPQAKPEEQENDPVGGA GYKPQMHLPINSTVEDIAAEEDLDKTAGYRP QANVNTWNLVSPDSPRSIDSNSEIVSFGSPCSI NSRQFLIPPKDEDSPKSNGGGWSFTNFFQNKP ND
1341	2691	A	10392	1		MLPPKHLSATKPKKSWAPNLYELDSDLTKEP DVIIGEGPTDSEFFHQRFRNLIYVEFVGPRKTL IKLRNLCLDWLQPETRTKEEIIELLVLEQYLTII PEKLKPWVRAKKPENCEKLVTLLENYKEMY QPEGESLHGVLVVSAGLRCPLGLSASTLLTW SGLDNSLSWAAVGMSCVLWDIELHHDFLGV ATKSVSTHAQGDAAQGLGGTIVRMWARDSN LATGVLLDDNNSDVTSDDDMTRNRESSPPH SVHSFSGDRDWDRRGRSRDTEPRDRWSHTR NPRSRMPPRDLSLPVVAKTSFEMDREDDRDS RAYESRSQDAESYQNVVDLAEDRKPHNTIQD NMENYRKLLSLGVQLAEDDGHSHMTQGHSS RSKRSAYPSTSRGLKTMPEAKKSTHRRGICED ESSHGVIMEKFIKDVSRSSKSGRARESSDRSQ RFPRMSDDNWKDISLNKRESVIQQRVYEGNA FRGGFRFNSTLVSRKRVLERKRYHFDTDGK GSIHDQKGCPRKKPFECGSEMRKAMSVSSLS SLSSPSFTESQPIDFGAMPYVCDECGRSFSVIS EFVEHQIMHTRENLYEYGESFIHSVAVSEVQK

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ŀ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine.
1	ĺ	ĺ	ĺ	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1			Í	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1		İ	ĺ	peptide	í ·	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
			1			SQVGGKRFECKDCGETFNKSAALAEHRKIHA
1		1	i		İ	RGYLVECKNQECEEAFMPSPTFSELQKIYGK
1			ļ		5	DKFYECRVCKETFLHSSALIEHQKIHFGDDKD
1	ĺ	[NEREHEREREREGETFRPSPALNEFOKMYG
						KEKMYECKVCGETFLHSSSLKEHQKIHTRGN
1			ĺ			PFENKGKVCEETFIPGQSLKRRQKTYNKEKLC
i						DFTDGRDAFMQSSELSEHQKIHSRKNLFEGR
						GYEKSVIHSGPFTESQKSHTITRPLESDEDEKA
ł		ļ			}	FTISSNPYENQKIPTKENVYEAKSYERSVIHSL
ŀ						ASVEAQKSHSVAGPSKPKVMAESTIQSFDAIN
ì		1				HQRVRAGGNTSEGREYSRSVIHSLVASKPPRS
						HNGNELVESNEKGESSIYISDLNDKROKIPAR
		}				ENPCEGGSKNRNYEDSVIQSVFRAKPQKSVP
]						GEGSGEFKKDGEFSVPSSNVREYQKARAKKK
						YIEHRSNETSVIHSLPFGEQTFRPRGMLYECQ
						ECGECFAHSSDLTEHQKIHDREKPSGSRNYE
						WSVIRSLAPTDPQTSYAQEQYAKEQARNKCK
i					'	DFRQFFATSEDLNTNQKIYDQEKSHGEESQGE
-						NTDGEETHSEETHGQETIEDPVIQGSDMEDPQ
						KDDPDDKIYECEDCGLGFVDLTDLTDHQKVH
						SRKCLVDSREYTHSVIHTHSISEYQRDYTGEQ
						LYECPKCGESFIHSSFLFEHQRIHEQDQLYSM
						KGCDDGFIALLPMKPRRNRAAERNPALAGSA
						IRCLLCGQGFIHSSALNEHMRLHREDDLLEQS
						QMAEEAIIPGLALTEFQRSQTEERLFECAVCG
						ESFVNPAELADHVTVHKNEPYEYGSSYTHTS
1						FLTEPLKGAIPFYECKDCGKSFIHSTVLTKHKE
						LHLEEEEEDEAAAAAAAAAQEVEANVHVPQ
				' i	1	VVLRIQGLNVEAAEPEVEAAEPEV
			,	1		EAAEPNGEAEGPDGEAAEPIGEAGQPNGEAE
			l i			QPNGDADEPDGAGIEDPEERAEEPEGKAEEPE
			Ì	· .		GDADEPDGVGIEDPEEGEDQEIQVEEPYYDC
				}		HECTETFTSSTAFSEHLKTHASMIIFEPANAFG
				Ì	ĺ	ECSGYIERASTSTGGANQADEKYFKCDVCGQ
						LFNDHLSLARHQNTHTG
1342	2692	A	10393	2	1350	GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA
						ASRRSPAARPPVPAPPALPRGRPGTEGSTSLS
			.			APAVLVVAVAVVVVVVSAVAWAMANYIHV
			j	ļ	,	PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ
						HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN
1						TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL
						QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH
						VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL
						WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI
[ĺ	İ	1		LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII
j l		ļ				YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE
		1	1	ſ		FAGVSDVDYSLYPDRELQSQWLRAYLEAYK
	Ì		}	İ		EFKGFGTEVTEKEVEILFIQVNQFALASHFFW
	-			Ĭ		GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM
L						KPEVTALKVPE
1343	2693	Α	10394	102	839	PEAQTSAVLAREKGHLPTMRHEAPMQMASA
	l	ł	ł	ł		QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK
		ļ				TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN
1		1	ł	ł	}	EKRIKLQIWDTAGQERYRTTTTAYYRGAMGFI
		l				LMYDITNEESFNAVQDWSTQIKTYSWDNAQ
			ĺ	ľ		VILVGNKCDMEDERVISTERGQHLGEQLGFE
		1	ļ			FFETSAKDNINVKQTFERLVDIICDKMSESLET
				i	_	DPAITAAKQNTRLKETPPPPQPNCAC
1344	2694	Α	10395	2	4136	DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS
				İ	1	LQQRTPAEMSPVLHFYVRPSGHEGAASGHTR
	,					

SEQ ID NO: of nucl- eotidc seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valinc, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion RKLQGKLPELQGVETELCYNVNWTAEALPSA EETKKLMWLFGCPLLLDDVARESWLLPGSN DLLLEVGPRLNFSTPTSTNIVSVCRATGLGPV DRVETTRRYRLSFAHPPSAEVEAIALATLHDR MTEQHFPHPIQSFSPESMPEPLNGPINILGEGR LALEKANQELGLALDS WDLDFYTKRFQELQR NPSTVEAFDLAQSNSEHSRHWFFKGQLHVDG QKLVHSLFESIMSTQESSNPNNVLKFCDNSSA IQGKEVRFLRPEDPTRPSRFQQQQGLRHVVFT AETHNFPTGVCPFSGATTGTGGRIRDVQCTG RGAHVVAGTAGYCFGNLHIPGYNLPWEDLSF QYPGNFARPLEVAIEASNGASDYGNKFGEPV LAGFARSLGLQLPDGQRREWIKPIMFSGGIGS MEADHISKEAPEPGMEVVKVGGPVYRIGVGG GAASSVQVQGDNTSDLDFGAVQRGDPEMEQ KMNRVIRACVEAPKGNPICSLHDQGAGGNG NVLKELSDPAGAIIYTSRFQLGDPTLNALEIW GAEYQESNALLLRSPNRDFLTHVSARERCPA CFVGTITGDRRIVLVDDRECPVRRNGQGDAP PTPPPTPVDLELEWVLGKMPRKEFFLQRKPP MLQPLALPPGLSVHQALERVLRLPAVASKRY LTNKVDRSVGGLVAQQCVGPLQTPLADVA VVALSHEELIGAATALGEQPVKSLLDPKVAA RLAVAEALTNLVFALVTDLRDVKCSGNWM WAAKLPGEGAALADACEAMVAVMAALGVA VDGGKDSLSMAARVGTETVRAPGSLVISAYA VCPDITATVTPDLKHPEGRGHLLYVALSPGQ HRLGGTALAQCFSQLGEHPPDLDLPENLVRA FSITQGLLKDRLLCSGHDVSDGGLVTCLLEM AFAGNCGLQVDVPVPRVDVLSVLFAEEPGLV LEVQEPDLAQVLKRYRDAGLHCLELGHTGE
						AGPHAMVRVSVNGAVVLEEPVGELRALWEE TSFQLDRLQAEPRCVAEEERGLRERMGPSYC LPPTFPKASVPREPGGPSPRVAILREEGSNGDR EMADAFHLAGFEVWDVTMQDLCSGAIGLDT FRGVAFVGGFSYADVLGSAKGWAAAVTFHP RAGAELRFFKRPDTFSLGVCNGCQLLALLG WVGGDPNEDAAEMGPDSQPARPGLLRHNL SGRYESRWASVRVGPGPALMLRGMEGAVLP VWSAHGEGYVAFSSPELQAQIEARGLAPLHW ADDDGNPTEQYPLNPNGSPGGVAGICSCDGR HLAVMPHPERAVRPWQWAWRPPPFDTLTTS PWLQLFINARNWTLEGSC
1345	2695	А	10396	65	642	GVRGFWAGTMASRAGPRAAGTDGSDFQHRE RVAMHYQMSVTLKYEIKKLIYVHLVIWLLLV AKMSVGHLRLLSHDQVAMPYQWEYPYLLSI LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPLI YGSMEMFPAAQQLYRHGKAYRFLFGFSAVSI MYLVLVLAVQVHAWQLYYSKKLLDSWFTST QEKKHK
1346	2696	A	10398	1	718	DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRGT EIGVLAKAFIDQGKLIPDDVMTRLALHELKNL TQYSWLLDGFPRTLPQAEALDRAYQIDTVINL NVPFEVIKQRLTARWIHPASGRVYNIEFNPPK TVGIDDLTGEPLIQREDDKPETVIKRLKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTP
1347	2697	Α	10402	153	1969	KHRQENNALDMAPEIHMTGPMCLIENTNGEL VANPEALKILSAITQPVVVVAIVGLYRTGKSY

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ļ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1		!		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		j	ļ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	,	/=possible nucleotide deletion, \=possible
		ļ	j	sequence	ļ	nucleotide insertion
						LMNKLAGKNKGFSLGSTVKSHTKGIWMWCV
		ļ]	J		PHPKKPEHTLVLLDTEGLGDVKKGDNQNDS
				ļ		WIFTLAVLLSSTLVYNSMGTINQQAMDQLYY
			1			VTELTHRIRSKSSPDENENEDSADFVSFFPDFV
Í			1			WTLRDFSLDLEADGQPLTPDEYLEYSLKLTQ
						GTSQKDKNFNLPRLCIRKFFPKKKCFVFDLPI
			Ì	ĺ		HRRKLAQLEKLQDEELDPEFVQQVADFCSYI
						FSNSKTKTLSGGIKVNGPRLESLVLTYINAISR
1			ł	ł		GDLPCMENAVLALAQIENSAAVQKAIAHYD
			ļ			QQMGQKVQLPAETLQELLDLHRVSEREATEV
}		l	l			YMKNSFKDVDHLFQKKLAAQLDKKRDDFCK
			Ì			QNQEASSDRCSALLQVIFSPLEEEVKAGIYSK
)			PGGYCLFIQKLQDLEKKYYEEPRKGIQAEEIL
						QTYLKSKESVTDAILQTDQILTEKEKEIEVEC
		ļ				VKAESAQASAKMVEEMQIKYQQMMEEKEKS
						YOEHVKOLTEKMERERAQLLEEQEKTLTSKL
		1		ĺ		QEQARVLKERCQGESTQLQNEIQKLQKTLKK
						KTKRYMSHKLKI
1348	2698	A	10404	5	892	TQLPAPLSGVLSRLQLGSGAPLLTWVQETAG
İ			·			VAGGAPRRRTPVTMWRLLARASAPLLRVPLS
		1	ľ			DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL
		1				RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT
]			}			EGNFAILALGGGYLHWGHFEMMRLTINRSM
1						DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK
1 .]		j		GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG
				ľ		FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE
						RERNNQNPWTFERIATANMLGIRKVLSPYDL
						THKGKYWGKFYMPKRV
1349	2699	Α	10409	59	1184	LRRNCSALGGLFQTIISDMKGSYPVWEDFINK
						AGKLQSQLRTTVVAAAAFLDAFQKVADMAT
						NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS
						ALIDCLINPLQEQMEEWKKVANQLDKDHAK
						EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ
ľ					<u> </u>	PQLDSALQDVNDKYLLLEETEKQAVRKALIE
						ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE
			,			DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS
 '						YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS
						GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH
]]			DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE
						PDPNGGGPTTASGPPAAAEEAQRPRSM
1350	2700	À	10410	511	958	AGRGGPGKPVSWSSGPGSPGQTQRRSWVKST
						RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW
1 1			1			AFNLYDLNKDGCITKEEMLDIMKSIYDMMG
1 1						
1						KYTYPALREEAPREHVESFFQKMDRNKDGV VTIEEFIESCQKDENIMRSMQLFDNVI

WHAT IS CLAIMED IS:

1. An isolated polynucleotide comprising a nucleotide sequence selected from the group consisting of SEQ ID NO: 1-1350, a mature protein coding portion of SEQ ID NO: 1-1350, an active domain of SEQ ID NO: 1-1350, and complementary sequences thereof.

- 2. An isolated polynucleotide encoding a polypeptide with biological activity, wherein said polynucleotide hybridizes to the polynucleotide of claim 1 under stringent hybridization conditions.
- 3. An isolated polynucleotide encoding a polypeptide with biological activity, wherein said polynucleotide has greater than about 90% sequence identity with the polynucleotide of claim 1.
- 4. The polynucleotide of claim 1 wherein said polynucleotide is DNA.
- 5. An isolated polynucleotide of claim 1 wherein said polynucleotide comprises the complementary sequences.
- 6. A vector comprising the polynucleotide of claim 1.
- 7. An expression vector comprising the polynucleotide of claim 1.
- 8. A host cell genetically engineered to comprise the polynucleotide of claim 1.
- 9. A host cell genetically engineered to comprise the polynucleotide of claim 1 operatively associated with a regulatory sequence that modulates expression of the polynucleotide in the host cell.
- 10. An isolated polypeptide, wherein the polypeptide is selected from the group consisting of:
 - (a) a polypeptide encoded by any one of the polynucleotides of claim 1; and
 - (b) a polypeptide encoded by a polynucleotide hybridizing under stringent conditions with any one of SEQ ID NO:1-1350.
- 11. A composition comprising the polypeptide of claim 10 and a carrier.
- 12. An antibody directed against the polypeptide of claim 10.

13. A method for detecting the polynucleotide of claim 1 in a sample, comprising:

- a) contacting the sample with a compound that binds to and forms a complex with the polynucleotide of claim 1 for a period sufficient to form the complex; and
- b) detecting the complex, so that if a complex is detected, the polynucleotide of claim 1 is detected.
- 14. A method for detecting the polynucleotide of claim 1 in a sample, comprising:
- a) contacting the sample under stringent hybridization conditions with nucleic acid primers that anneal to the polynucleotide of claim 1 under such conditions;
- b) amplifying a product comprising at least a portion of the polynucleotide of claim 1; and
- c) detecting said product and thereby the polynucleotide of claim 1 in the sample.
- 15. The method of claim 14, wherein the polynucleotide is an RNA molecule and the method further comprises reverse transcribing an annealed RNA molecule into a cDNA polynucleotide.
- 16. A method for detecting the polypeptide of claim 10 in a sample, comprising:
- a) contacting the sample with a compound that binds to and forms a complex with the polypeptide under conditions and for a period sufficient to form the complex; and
- b) detecting formation of the complex, so that if a complex formation is detected, the polypeptide of claim 10 is detected.
- 17. A method for identifying a compound that binds to the polypeptide of claim 10, comprising:
- a) contacting the compound with the polypeptide of claim 10 under conditions sufficient to form a polypeptide/compound complex; and
- b) detecting the complex, so that if the polypeptide/compound complex is detected, a compound that binds to the polypeptide of claim 10 is identified.
- 18. A method for identifying a compound that binds to the polypeptide of claim 10, comprising:

a) contacting the compound with the polypeptide of claim 10, in a cell, under conditions sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a reporter gene sequence in the cell; and

- b) detecting the complex by detecting reporter gene sequence expression, so that if the polypeptide/compound complex is detected, a compound that binds to the polypeptide of claim 10 is identified.
- 19. A method of producing the polypeptide of claim 10, comprising,
- a) culturing a host cell comprising a polynucleotide sequence selected from the group consisting of a polynucleotide sequence of SEQ ID NO: 1-1350, a mature protein coding portion of SEQ ID NO: 1-1350, an active domain of SEQ ID NO: 1-1350, complementary sequences thereof and a polynucleotide sequence hybridizing under stringent conditions to SEQ ID NO: 1-1350, under conditions sufficient to express the polypeptide in said cell; and
 - b) isolating the polypeptide from the cell culture or cells of step (a).
- 20. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 1351-2700, the mature protein portion thereof, or the active domain thereof.
- 21. The polypeptide of claim 20 wherein the polypeptide is provided on a polypeptide array.
- 22. A collection of polynucleotides, wherein the collection comprises the sequence information of at least one of SEQ ID NO: 1-1350.
- 23. The collection of claim 22, wherein the collection is provided on a nucleic acid array.
- 24. The collection of claim 23, wherein the array detects full-matches to any one of the polynucleotides in the collection.
- 25. The collection of claim 23, wherein the array detects mismatches to any one of the polynucleotides in the collection.
- 26. The collection of claim 22, wherein the collection is provided in a computer-readable format.

27. A method of treatment comprising administering to a mammalian subject in need thereof a therapeutic amount of a composition comprising a polypeptide of claim 10 or 20 and a pharmaceutically acceptable carrier.

28. A method of treatment comprising administering to a mammalian subject in need thereof a therapeutic amount of a composition comprising an antibody that specifically binds to a polypeptide of claim 10 or 20 and a pharmaceutically acceptable carrier.

Pages 340 to 1963 of this application contain amino acid sequence listings. They can be obtained at the address given below.

Les pages 340 to 1963 de cette demande contiennent des listages des séquences d'acides aminés. Elles peuvent être obtenues à l'adresse indiquée ci-dessous.

World Intellectual Property Organization 34, chemin des Colombettes CH-1211 Genève 20